GENE EXPRESSION MODULATED IN GASTROINTESTINAL INFLAMMATION

REFERENCE TO RELATED APPLICATION

This application claims the benefit of U.S. Provisional Application S.N. 60/138,487, filed June 10, 1999, which is incorporated herein by reference.

BACKGROUND OF THE INVENTION

Chronic inflammatory bowel diseases (IBD) are as a persistent problem in medicine. The two major types of IBD are Crohn's disease (CD), which can affect the whole digestive tract from mouth to anus, and ulcerative colitis (UC), which affects only the large intestine.

Although CD and UC are both characterized by massive gut damage arising from intestinal inflammation, these diseases are quite distinct. Macroscopically, the best distinguishing features of these diseases are that UC is 1) a continuous inflammatory disorder, 2) primarily restricted to the mucosa of the colon and rectum, 3) usually primarily vascular, and 4) usually associated with a striking shortening of the colon. In contrast, inflammation in CD is nearly always discontinuous and may present anywhere in the large or small intestine. With CD the mucosa usually appears 'cobblestoned' and there is fissuring and a thickening of the bowel wall with associated stenosis of the lumen and strictures.

Microscopically, UC is essentially a superficial inflammation of the mucous membrane and even in chronic disease, the muscularis propria and serosa are free of inflammatory infiltration. In contrast, CD presents microscopically as a transmural inflammation, spreading through the mucosa and submucosa into the muscle layers. In UC, crypt abscesses and destruction of the epithelium are common, whereas in CD, a valuable diagnostic feature is the presence of sarcoid type granulomas.

The histological features of both of these diseases suggest that there is a dysregulation of the lymphoid tissue in IBD. Whether the immunological dysregulation is a primary cause

of IBD or whether the inflammatory response is secondary to another mucosal insult, such as a breach in the epithelial barrier, remains unclear.

Despite advances in recent years, the precise etiology and pathogenesis of CD and UC remain undefined. In order to try and better understand the mechanistic basis of IBD, much effort has been directed towards the discovery and development of different animal models of inflammatory bowel disease. One of the best characterized systems for studying intestinal inflammation is the mouse model of dextran sodium sulfate (DSS)-induced colitis (I. Okayasu et al, Gastroenterology 98:694702, 1990; H.S. Cooper et al, Laboratory Investigation 69:238-249, 1993; L.A. Dielman et al, Gastroenterology 107:1643-1652, 1994; C.O. Elson et al, Gastroenterology 109:1344-1367, 1995).

DSS is thought to cause breaches in the epithelial tight junctions (J. Ni et al, Gut 39:234-241, 1996), thus permitting the huge antigenic load in the gut lumen to come into direct contact with the underlying tissues. As a result, a vigorous immune response directed against antigens located in the gut lumen is initiated in DSS-treated mice. The intestinal inflammation is primarily restricted to the large intestine.

The DSS-induced colitis model system is used to examine (i) the nature of the earliest response to epithelial damage, (ii) the mechanisms responsible for recruiting cells to the sites of inflammation, (iii) the nature of the protective immune response at the height of intestinal inflammation and (iv) the mechanisms that direct recovery and trigger the repair of damaged tissues. This model system can be used to examine the mechanisms of induction and recovery of IBD and should aid the identification of genes/proteins that may be able to modulate or prevent intestinal damage or stimulate recovery of the mucosa.

Given the diversity of factors that may contribute to these processes, a clear need is evident for the discovery, identification and elucidation of the roles of new proteins involved in the different stages of IBD.

Features of DSS-induced colitis

In the DSS-induced colitis model system, weight loss is apparent in mice beginning at

day 4. Histological analysis of the intestine reveals the presence of early patchy lesions identifiable by the loss of epithelial cells and goblet cells. By day 8, weight loss is fairly severe (approximately 20% reduction) and the mice appear moribund. Histologically, the gut epithelium is almost totally destroyed at this stage. There is evidence of a large mixed inflammatory cell infiltration into the lamina propria and submucosa. The inflammatory cell infiltrate appears to be composed primarily of T cells, B cells and granulocytes. By day 12, weight gain is apparent as the mice recover. At this later stage, crypt recovery and epithelial regeneration provide histological evidence of the beginning of repair processes.

DSS-induced colitis and human IBD

The primary event in DSS-induced colitis is epithelial cell damage, with inflammation associated with immune activation a secondary event. Whether epithelial cell damage or immune dysregulation is the triggering event in human IBD is unclear. The model of DSS-colitis is therefore useful in providing an opportunity to study the development and consequences of both processes.

SUMMARY OF THE INVENTION

Molecules have been identified that correspond to genes that are regulated by the DSS treatment. Such molecules are useful in therapeutic, prognostic and diagnostic applications in the treatment of IBD and other gut pathologies. The present invention provides novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polynucleotides and the polypeptides. Also provided are diagnostic methods for detecting disorders related to the polypeptides and the polynucleotides encoding them, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

The present invention provides novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polynucleotides and the polypeptides. Also provided are diagnostic methods for detecting disorders related to the polypeptides and the polynucleotides

encoding them, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

In one embodiment, the invention provides an isolated nucleic acid molecule comprising a polynucleotide chosen from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61 and SEQ ID NO:62. Another embodiment comprises an isolated nucleic acid molecule at least 95% identical to the isolated nucleic acid molecule of SEQ ID NO:1-62. A further embodiment comprises an isolated nucleic acid molecule at least ten bases in length that is hybridizable to the isolated nucleic acid molecule of SEQ ID NO:1-62 under stringent conditions.

In another embodiment, the invention provides an isolated polypeptide encoded by a polynucleotide chosen from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58,

SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61 and SEQ ID NO:62. In another embodiment, the invention provides an isolated nucleic acid molecule encoding the polypeptide of the present invention.

In a further embodiment, the invention provides a substantially pure isolated DNA molecule suitable for use as a probe for genes regulated in gastrointestinal inflammation, chosen from the group consisting of the DNA molecules identified in Table 1, having a 5' partial nucleotide sequence and length as described by their digital address, and having a characteristic regulation pattern in gastrointestinal inflammation.

The present invention also provides a system and method for detecting the presence of a gene regulated in gastrointestinal inflammation. In one embodiment, the present invention provides a kit for suitable for detecting the presence of a gene regulated in gastrointestinal inflammation, comprising at least one polynucleotide of the present invention, or fragment thereof having at least 10 contiguous bases, in an amount sufficient for at least one assay; label means; instructions for use; and suitable packaging material. In one embodiment, the polynucleotide is chosen from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61 and SEQ ID NO:62. Another embodiment comprises a polynucleotide at least 95% identical to the isolated nucleic acid molecule of SEQ ID NO:1-62. A further embodiment comprises a polynucleotide at least ten bases in length that is hybridizable to the isolated nucleic acid molecule of SEQ ID NO:1-62 under stringent conditions. In yet another embodiment, the polynucleotide is chosen from the group consisting of the DNA molecules identified in Table 1, having a 5' partial nucleotide

sequence and length as described by their digital address, and having a characteristic regulation pattern in gastrointestinal inflammation.

BRIEF DESCRIPTION OF THE DRAWINGS

These and other features, aspects, and advantages of the present invention will become better understood with reference to the following description, appended claims, and accompanying drawings where:

Figure 1 is a graphical representation of the results of TOGA analysis (TOtal Gene expression Analysis) using a 5' PCR primer with parsing bases CGCG, showing PCR products produced from mRNA extracted from colons isolated from untreated mice (Figure 1A, "0 Days"), mRNA extracted from colons isolated from mice receiving DSS for four days (Figure 1B, "4 Days"), mRNA extracted from colons isolated from mice receiving DSS for eight days (Figure 1C, "8 Days"), and mRNA extracted from colons isolated from mice receiving DSS for twelve days (Figure 1D, "12 Days"), where the vertical index line indicates a PCR product of about 458 b.p. that is up-regulated by DSS treatment, reaching a maximum at eight days, where the ordinate is in arbitrary units of fluorescence intensity and the abscissa is length of PCR product in nucleotides;

Figure 2 is a graphical representation of more detailed analysis of the 458 b.p. PCR product indicated in Figure 1; Figure 2A shows the PCR product obtained using an extended 5' primer as described in the text; Figure 2B shows the PCR products obtained using the original PCR primers, and in Figure 2C, the traces from Figure 2A and 2B are overlaid, demonstrating that the PCR product of the isolated and sequenced clone is the same length as the original PCR product, where the ordinate is in arbitrary units of fluorescence intensity and the abscissa is length of PCR product in nucleotides;

Figure 3 is a graphical representation of the results of Northern Blot analysis of clone IMX 2_46, SEQ ID NO: 10, where an agarose gel containing poly A enriched mRNA from the four experimental samples (0, 4, 8 or 12 days DSS treatment) as well as size standards was blotted after electrophoresis and probed with radiolabelled IMX 2_46, SEQ ID NO: 10, imaged using a phosphorimager and quantified. Quantitative results showing the relative

expression levels of the 1.6 kb transcript were: 0 day, 64; 4 days, 53; 8 days, 223; and 12 days, 269. The amount of RNA loaded on the gel was determined by probing for cyclophilin ("cyc").

Figure 4 is a graphical representation of the results of RT-PCR of clone IMX2_48.

Figure 5 is a graphical representation of the results of RT-PCR of clone IMX2_74.

Figure 6 is a graphical representation of the results of Northern blot analysis of clone IMX 2_17, SEQ ID NO: 3, where an agarose gel containing poly A enriched mRNA from the experimental samples and size standards was blotted after electrophoresis, probed, imaged using a phosphorimager and quantified. Figure 6A shows the results from C57BL/6 mice with DSS colitis 0, 4, 8, or 12 days after treatment, C57BL/6 mice with αCD3 ileitis 0, 6, 30 or 72 hours after treatment, as well as samples from large intestines from FVB, mdr knock-out mice without colitis and mdr knock-out mice with colitis. Figure 6B shows the results from Balb/c mice with DSS colitis 0, 4, 8, or 12 days after treatment, Balb/c mice treated with 0%, 5% and 8% DSS, and Balb/c mice with αCD3 ileitis 0, 6, 30 or 72 hours after treatment. The predicted transcript size for IMX2_17 CRP-ductin is 6.6 Kb; the actual transcript size found in this study was approximately 6.5 Kb.

Figure 7 is a graphical representation of the results of Northern blot analysis of clone IMX 2_22, SEQ ID NO: 4, where an agarose gel containing poly A enriched mRNA from the experimental samples and size standards was blotted after electrophoresis, probed, imaged using a phosphorimager and quantified. Figure 7A shows the results from C57BL/6 mice with DSS colitis 0, 4, 8, or 12 days after treatment, C57BL/6 mice with αCD3 ileitis 0, 6, 30 or 72 hours after treatment, as well as samples from large intestines from FVB, mdr knock-out mice without colitis, mdr knock-out mice with colitis and C57BL/6 spleen. Figure 7B shows the results from Balb/c mice with DSS colitis 0, 4, 8, or 12 days after treatment, Balb/c mice with αCD3 ileitis 0, 6, 30 or 72 hours after treatment, and C57BL/6 normal lymphoid tissue samples (MLM, PP, spleen and thymus). The predicted transcript size for IMX2_22 HPK1 is 2.7 Kb; the actual transcript size found in this study was approximately 2.8 Kb.

Figure 8 is a graphical representation of the results of Northern blot analysis of clone IMX 2_28, SEQ ID NO: 5, where an agarose gel containing poly A enriched mRNA from the experimental samples and size standards was blotted after electrophoresis, probed, imaged using a phosphorimager and quantified. Figure 8A shows the results from C57BL/6 mice with DSS colitis 0, 4, 8, or 12 days after treatment, C57BL/6 mice with αCD3 ileitis 0, 6, 30 or 72 hours after treatment, as well as samples from large intestines from FVB, mdr knock-out mice without colitis and mdr knock-out mice with colitis. Figure 8B shows the results from Balb/c mice with DSS colitis 0, 4, 8, or 12 days after treatment, Balb/c mice treated with 0%, 5% and 8% DSS, and Balb/c mice with αCD3 ileitis 0, 6, 30 or 72 hours after treatment. The predicted transcript size for IMX2_28 DRA is 2.6 Kb; the actual transcript size found in this study was approximately 3 Kb.

Figure 9 is a graphical representation of the results of Northern blot analysis of clone IMX 2_33, SEQ ID NO:21, where an agarose gel containing poly A enriched mRNA from the experimental samples and size standards was blotted after electrophoresis, probed, imaged using a phosphorimager and quantified. Figure 9A shows the results from C57BL/6 mice with DSS colitis 0, 4, 8, or 12 days after treatment, C57BL/6 mice with αCD3 ileitis 0, 6, 30 or 72 hours after treatment, as well as samples from large intestines from FVB, mdr knock-out mice without colitis and mdr knock-out mice with colitis. Figure 9B shows the results from Balb/c mice with DSS colitis 0, 4, 8, or 12 days after treatment, Balb/c mice treated with 0%, 5% and 8% DSS, and Balb/c mice with αCD3 ileitis 0, 6, 30 or 72 hours after treatment. The predicted transcript size for IMX2_33 SLPI is 1.1 Kb; the actual transcript size found in this study was 1.1 Kb.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by

the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:1-62. For example, the polynucleotide can contain all or part of the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:1-62, or the complement thereof, or the cDNA. "Stringent hybridization conditions" refers to an overnight incubation at 42° C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt

conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH_2PO_4; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 ug/ml salmon sperm blocking DNA;

followed by washes at 50°C with 1x SSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5x SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can

be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formulation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS – STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention,

including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.).

The translated amino acid sequence, beginning with the methionine, is identified although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by the translation of these alternative open reading frames are specifically contemplated by the present invention.

SEQ ID NO:1-62 and the translations of SEQ ID NO:1-62 are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from the translations of SEQ ID NO:1-62 may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

The present invention also relates to the genes corresponding to SEQ ID NO:1-62, and translations of SEQ ID NO:1-62. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologues. Species homologues may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below). It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

Signal Sequences

Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch, Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, Nucleic Acids Res. 14:4683-4690 (1986) uses the information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1 indicates the amino terminus of the secreted protein. Therefore, from a deduced amino acid sequence, a signal sequence and mature sequence can be identified.

Polynucleotide and Polypeptide Variants

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

"Identity" per se has an art-recognized meaning and can be calculated using published techniques. (See, e.g.: (COMPUTATIONAL MOLECULAR BIOLOGY, Lesk, A.M., ed., Oxford University Press, New York, (1988); BIOCOMPUTING: INFORMATICS AND GENOME PROJECTS, Smith, D.W., ed., Academic Press, New York, (1993); COMPUTER ANALYSIS OF SEQUENCE DATA, PART I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, (1994); SEQUENCE ANALYSIS IN MOLECULAR BIOLOGY, von Heinje, G., Academic Press, (1987); and SEQUENCE ANALYSIS PRIMER, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, (1991).) While there exists a number of methods to measure identity between two polynucleotide or polypeptide sequences, the term "identity" is well known to skilled artisans. (Carillo, H., and Lipton, D., SIAM J Applied Math 48:1073 (1988).) Methods commonly employed to determine identity or similarity between two sequences include, but are not limited to, those disclosed in "Guide to Huge Computers," Martin J. Bishop, ed., Academic Press, San Diego, (1994), and Carillo, H., and Lipton, D., SIAM J Applied Math 48:1073 (1988). Methods for aligning polynucleotides or polypeptides are codified in computer programs, including the GCG program package (Devereux, J., et al., Nucleic Acids Research (1984) 12(l):387 (1984)), BLASTP, BLASTN, FASTA (Atschul, S.F. et al., J. Molec. Biol. 215:403 (1990), Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711 (using the local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2:482-489 (1981).)

When using any of the sequence alignment programs to determine whether a particular sequence is, for instance, 95% identical to a reference sequence, the parameters are set so that the percentage of identity is calculated over the full length of the reference polynucleotide and that gaps in identity of up to 5% of the total number of nucleotides in the reference polynucleotide are allowed.

A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global

sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. 6:237-245 (1990).) The term "sequence" includes nucleotide and amino acid sequences. In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB search of a DNA sequence to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=l, Joining Penalty=30, Randomization Group Length=0, and Cutoff Score=l, Gap Penalty=5, Gap Size Penalty 0.05, and Window Size=500 or query sequence length in nucleotide bases, whichever is shorter. Preferred parameters employed to calculate percent identity and similarity of an amino acid alignment are: Matrix=PAM 150, k-tuple=2, Mismatch Penalty= 1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=l, Gap Penalty=5, Gap Size Penalty=0.05, and Window Size=500 or query sequence length in amino acid residues, whichever is shorter.

As an illustration, a polynucleotide having a nucleotide sequence of at least 95% "identity" to a sequence contained in SEQ ID NO:1-62 means that the polynucleotide is identical to a sequence contained in SEQ ID NO:1-62 or the cDNA except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the total length (not just within a given 100 nucleotide stretch). In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to SEQ ID NO:1-62, up to 5% of the nucleotides in the sequence contained in SEQ ID NO:1-62 or the cDNA can be deleted, inserted, or substituted with other nucleotides. These changes may occur anywhere throughout the polynucleotide.

Further embodiments of the present invention include polynucleotides having at least 80% identity, more preferably at least 90% identity, and most preferably at least 95%, 96%, 97%, 98% or 99% identity to a sequence contained in SEQ ID NO:1-62. Of course, due to the degeneracy of the genetic code, one of ordinary skill in the art will immediately recognize that a large number of the polynucleotides having at least 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity will encode a polypeptide identical to an amino acid sequence contained in the translations of SEQ ID NO:1-62.

Similarly, by a polypeptide having an amino acid sequence having at least, for example, 95% "identity" to a reference polypeptide, is intended that the amino acid sequence of the polypeptide is identical to the reference polypeptide except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the total

length of the reference polypeptide. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

Further embodiments of the present invention include polypeptides having at least 80% identity, more preferably at least 85% identity, more preferably at least 90% identity, and most preferably at least 95%, 96%, 97%, 98% or 99% identity to an amino acid sequence contained in translations of SEQ ID NO:1-62. Preferably, the above polypeptides should exhibit at least one biological activity of the protein.

In a preferred embodiment, polypeptides of the present invention include polypeptides having at least 90% similarity, more preferably at least 95% similarity, and still more preferably at least 96%, 97%, 98%, or 99% similarity to an amino acid sequence contained in translations of SEQ ID NO:1-62.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host such as *E. coli*).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at

either the polynucleotide and/or polypeptide level. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988 (1993) reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-l a. They used random mutagenesis to generate over 3,500 individual IL- l a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See Gayle et al., (1993), Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on

activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp,
and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii)

fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

Polynucleotide and Polypeptide Fragments

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in that shown in SEQ ID NO:1-62. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in that shown in SEQ ID NO:1-62. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, and more nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, to the end of SEQ ID NO:1-62. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity.

In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in the translations of SEQ ID NO:1-62. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino

acid number 1-20, 21-40, 41-60, or 61 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50 or 60, amino acids in length. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of the translations of SEQ ID NO:1-62 falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Epitopes & Antibodies

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is

defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998-4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')2 fragments) which are capable of specifically binding to protein. Fab and F(ab')2 fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

Additional embodiments include chimeric antibodies, e.g., humanized versions of murine monoclonal antibodies. Such humanized antibodies may be prepared by known techniques, and offer the advantage of reduced immunogenicity when the antibodies are

administered to humans. In one embodiment, a humanized monoclonal antibody comprises the variable region of a murine antibody (or just the antigen binding site thereof) and a constant region derived from a human antibody. Alternatively, a humanized antibody fragment may comprise the antigen binding site of a murine monoclonal antibody and a variable region fragment (lacking the antigen-binding site) derived from a human antibody. Procedures for the production of chimeric and further engineered monoclonal antibodies include those described in Riechmann et al. (*Nature 332*:323, 1988), Liu et al. (*PNAS 84*:3439, 1987), Larrick et al. (*Bio/Technology 7*:934, 1989), and Winter and Harris (*TIPS 14*:139, May, 1993).

One method for producing an antibody comprises immunizing a non-human animal, such as a transgenic mouse, with a polypeptide translated from a nucleotide sequence chosen from SEQ ID NO:1-62, whereby antibodies directed against the polypeptide translated from a nucleotide sequence chosen from SEQ ID NO:1-62 are generated in said animal. Procedures have been developed for generating human antibodies in non-human animals. The antibodies may be partially human, or preferably completely human. Non-human animals (such as transgenic mice) into which genetic material encoding one or more human immunoglobulin chains has been introduced may be employed. Such transgenic mice may be genetically altered in a variety of ways. The genetic manipulation may result in human immunoglobulin polypeptide chains replacing endogenous immunoglobulin chains in at least some (preferably virtually all) antibodies produced by the animal upon immunization. Antibodies produced by immunizing transgenic animals with a polypeptide translated from a nucleotide sequence chosen from SEQ ID NO:1-62 are provided herein.

Mice in which one or more endogenous immunoglobulin genes are inactivated by various means have been prepared. Human immunoglobulin genes have been introduced into the mice to replace the inactivated mouse genes. Antibodies produced in the animals incorporate human immunoglobulin polypeptide chains encoded by the human genetic material introduced into the animal. Examples of techniques for production and use of such transgenic animals are described in U.S. Patents 5,814,318, 5,569,825, and 5,545,806, which are incorporated by reference herein.

Monoclonal antibodies may be produced by conventional procedures, e.g., by immortalizing spleen cells harvested from the transgenic animal after completion of the immunization schedule. The spleen cells may be fused with myeloma cells to produce hybridomas, by conventional procedures.

A method for producing a hybridoma cell line comprises immunizing such a transgenic animal with a immunogen comprising at least seven contiguous amino acid residues of a polypeptide translated from a nucleotide sequence chosen from SEQ ID NO:1-62; harvesting spleen cells from the immunized animal; fusing the harvested spleen cells to a myeloma cell line, thereby generating hybridoma cells; and identifying a hybridoma cell line that produces a monoclonal antibody that binds a polypeptide translated from a nucleotide sequence chosen from SEQ ID NO:1-62. Such hybridoma cell lines, and monoclonal antibodies produced therefrom, are encompassed by the present invention. Monoclonal antibodies secreted by the hybridoma cell line are purified by conventional techniques.

Antibodies may be employed in an *in vitro* procedure, or administered *in vivo* to inhibit biological activity induced by a polypeptide translated from a nucleotide sequence chosen from SEQ ID NO:1-62. Disorders caused or exacerbated (directly or indirectly) by the interaction of such polypeptides of the present invention with cell surface receptors thus may be treated. A therapeutic method involves *in vivo* administration of a blocking antibody to a mammal in an amount effective for reducing a biological activity induced by a polypeptide translated from a nucleotide sequence chosen from SEQ ID NO:1-62.

Also provided herein are conjugates comprising a detectable (e.g., diagnostic) or therapeutic agent, attached to an antibody directed against a polypeptide translated from a nucleotide sequence chosen from SEQ ID NO:1-62. Examples of such agents are well known, and include but are not limited to diagnostic radionuclides, therapeutic radionuclides, and cytotoxic drugs. The conjugates find use in *in vitro* or *in vivo* procedures.

Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover,

because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-0 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0 232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for

immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

Vectors, Host Cells, and Protein Production

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome

binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, 293, Bowes melanoma cells and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, PNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:1-62. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:1-62 will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be

achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library).) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined. The polynucleotides of SEQ ID NO:1-62 can be used for this analysis of individual humans.

First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from

several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen,

making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class H HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

Uses of the Polypeptides

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell . Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (^{125}I, ^{121}I), carbon (^{14}C), sulfur (^{35}S), tritium (^{3}H); indium (^{112}In), and technetium (^{99m}Tc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for *in vivo* imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, ^{131}I, ^{112}In, ^{99m}Tc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of ^{99m}Tc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in

Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

Biological Activities

The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the

diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

Immune Activitiv

A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

A polynucleotide or polypeptide of the present invention may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitis, and autoimmune inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions,

including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

Hyperproliferative Disorders

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenstron's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

Infectious Disease

A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiollitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsielia, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Menigococcal), Pasteurellacea Infections

(e.g., Actinobacillus, Heamophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

Regeneration

A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease

(e.g. osteoporosis, osteocarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vascular (including vascular endothelium), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stroke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease,
Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

Chemotaxis

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotaxic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotaxic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotaxic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors),or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2), Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled

competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity, and (c) determining if a biological activity of the polypeptide has been altered.

Other Activities

A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin,

percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, circadian rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by activin or inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

Other Preferred Embodiments

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 80%, preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:1-62.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:1-62 in the range of positions beginning with the nucleotide at about the position of the 5' nucleotide of the clone sequence and ending with the nucleotide at about the position of the 3' nucleotide of the clone sequence.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:1-62 in the range of positions beginning with the nucleotide at about the position of the 5' nucleotide of the start codon and ending with the nucleotide at about the position of the 3' nucleotide of the clone sequence as defined for SEQ ID NO:1-62.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:1-62 in the range of positions beginning with the nucleotide at about the position of the 5' nucleotide of the first

amino acid of the signal peptide and ending with the nucleotide at about the position of the 3' nucleotide of the clone sequence as defined for SEQ ID NO:1-62.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:1-62.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:1-62.

A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:1-62 beginning with the nucleotide at about the position of the 5' nucleotide of the first amino acid of the signal peptide and ending with the nucleotide at about the position of the 3' nucleotide of the clone sequence as defined for SEQ ID NO:1-62.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:1-62.

Also preferred is an isolated nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:1-62, which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:1-62.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of a nucleotide sequence of SEQ ID NO:1-62.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:1-62. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in an amino acid sequence translated from SEQ ID NO:1-62.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in acids in an amino acid sequence translated from SEQ ID NO:1-62, in the range of positions beginning with the residue at about the position of the first amino acid of the secreted portion and ending with the residue at about the last amino acid of the open reading frame.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in an amino acid sequence translated from SEQ ID NO:1-62.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in an amino acid sequence translated from SEQ ID NO:1-62.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to acids in an amino acid sequence translated from SEQ ID NO:1-62.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of amino acid sequences translated from SEQ ID NO:1-62, which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of amino acid sequences translated from SEQ ID NO:1-62.

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of amino acid sequences translated from SEQ ID NO:1-62.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of amino acid sequences translated from SEQ ID NO:1-62.

In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of amino acid sequences translated from SEQ ID NO:1-62.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encodes a polypeptide comprising an amino acid sequence selected from the group consisting of amino acid sequences translated from SEQ ID NO:1-62.

Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of amino acid sequences translated from SEQ ID NO:1-62. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

EXAMPLE 1

Identification and Characterization of Polynucleotides Regulated by *in vivo* DSS Treatment

For the induction of colitis, DSS (MW 40,000) is dissolved in the drinking water and given to mice ad libitum for a period of 7 days. The DSS water is then replaced with regular drinking water. Distinct mouse strains demonstrate differential susceptibility to DSS-colitis.

Three percent DSS is sufficient for the induction of acute colitis in C57BL/6 mice, whereas 5% DSS is required for the induction of acute colitis in BALB/c mice.

For the purpose of this study, samples of colon were harvested from cohorts of C57BL/6 mice treated with 3% DSS. Tissues were harvested at day 0, day 4, day 8 or day 12 - a schedule designed to encompass the full range of induction of intestinal damage through recovery.

Damage to colonic tissue is a hallmark of IBD. The mouse model of DSS-induced colitis gives rise to damaged colonic tissue and as such it is one of the more useful models for studying IBD.

Mice were treated with DSS in their drinking water as described above. At various intervals during DSS-treatment, cohorts of mice were sacrificed and the colons removed by dissection. Freshly dissected colonic tissue was immediately placed in GT buffer (4.5M guanidinium isothiocyanate, 50mM sodium citrate, 0.5%w/v sodium sarcosyl, 2% 2-beta-mercaptoethanol) and homogenized. Homogenized lysates were spun briefly to remove large debris before being layered onto a CsCl gradient. RNA was extracted using conventional methods and was subsequently used for TOGA analysis.

Features of DSS-induced colitis

Weight loss is apparent in mice beginning at day 4. Histological analysis of the intestine reveals the presence of early patchy lesions identifiable by the loss of epithelial cells and goblet cells. By day 8, weight loss is fairly severe (approximately 20% reduction) and the mice appear moribund. Histologically, the gut epithelium is almost totally destroyed at this stage. There is evidence of a large mixed inflammatory cell infiltration into the lamina propria and submucosa. The inflammatory cell infiltrate appears to be composed primarily of T cells, B cells and granulocytes. By day 12, weight gain is apparent as the mice recover. At this later stage, crypt recovery and epithelial regeneration provide histological evidence of the beginning of repair processes.

Isolated RNA was analyzed using a method of simultaneous sequence-specific identification of mRNAs known as TOGA (TOtal Gene expression Analysis) described in Sutcliffe, J.G., et al *Proc Natl Acad Sci U S A* 2000 Feb 29; 97(5):1976-1981, International published application PCT/US99/23655, U.S. Patent No. 5,459,037, U.S. Patent No. 5,807,680, and U.S. Patent No. 6,030,784, hereby incorporated herein by reference. Preferably, prior to the application of the TOGA method or other methods, the isolated RNA

was enriched to form a starting polyA-containing mRNA population by methods known in the art. In such a preferred embodiment, the TOGA method further comprises an additional Polymerase Chain Reaction ("PCR") step performed using one of four 5' PCR primers and cDNA templates prepared from a population of antisense complimentary RNA ("cRNAs"). A final PCR step using one of a possible 256 5' PCR primers and a universal 3' PCR primer produced as PCR products, cDNA fragments that corresponded to a 3'-region of the starting mRNA population. The produced PCR products were then identified by a) the sequence of at least the 5' seven base pairs, preferably the sequence of the entire fragment, and b) the length of the fragment. These two parameters, sequence and fragment length, were used to compare the obtained PCR products to a database of known polynucleotide sequences.

The method yields Digital Sequence Tags (DSTs), that is, polynucleotides that are expressed sequence tags (ESTs) of the 3' end of mRNAs. DSTs that showed changes in relative levels following DSS treatment were selected for further study. The intensities of the laser-induced fluorescence of the labeled PCR products were compared across sample isolated at different time intervals after treatment.

In general, double-stranded cDNA is generated from poly(A)-enriched cytoplasmic RNA extracted from the tissue samples of interest using an equimolar mixture of all 48 5'-biotinylated anchor primers of a set to initiate reverse transcription. One such suitable set is G-A-A-T-T-C-A-A-C-T-G-G-A-A-G-C-G-G-C-C-C-G-C-A-G-G-A-A-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-V-N-N (SEQ ID NO: 63), where V is A, C or G and N is A, C, G or T. One member of this mixture of 48 anchor primers initiates synthesis at a fixed position at the 3' end of all copies of each mRNA species in the sample, thereby defining a 3' endpoint for each species, resulting in biotinylated double stranded cDNA.

Each biotinylated double stranded cDNA sample was cleaved with the restriction endonuclease MspI, which recognizes the sequence CCGG. The 3' fragments of cDNA were then isolated by capture of the biotinylated cDNA fragments on a streptavidin-coated substrate. Suitable streptavidin-coated substrates include microtitre plates, PCR tubes, polystyrene beads, paramagnetic polymer beads and paramagnetic porous glass particles. A preferred streptavidin-coated substrate is a suspension of paramagnetic polymer beads (Dynal, Inc., Lake Success, NY).

After washing the streptavidin-coated substrate and captured biotinylated cDNA fragments, the cDNA fragment product was released by digestion with NotI, which cleaves at an 8-nucleotide sequence within the anchor primers but rarely within the mRNA-derived portion of the cDNAs. The 3' MspI-NotI fragments, which are of uniform length for each mRNA species, were directionally ligated into ClaI-, NotI-cleaved plasmid pBC SK^+ (Stratagene, La Jolla, CA) in an antisense orientation with respect to the vector's T3 promoter, and the product used to transform Escherichia coli SURE cells (Stratagene). The ligation regenerates the NotI site, but not the MspI site. Each library contained in excess of 5 x 10^5 recombinants to ensure a high likelihood that the 3' ends of all mRNAs with concentrations of 0.001% or greater were multiply represented. Plasmid preps (Qiagen) were made from the cDNA library of each sample under study.

An aliquot of each library was digested with MspI, which effects linearization by cleavage at several sites within the parent vector while leaving the 3' cDNA inserts and their flanking sequences, including the T3 promoter, intact. The product was incubated with T3 RNA polymerase (MEGAscript kit, Ambion) to generate antisense cRNA transcripts of the cloned inserts containing known vector sequences abutting the MspI and NotI sites from the original cDNAs.

At this stage, each of the cRNA preparations was processed in a three-step fashion. In step one, 250 ng of cRNA was converted to first-strand cDNA using the 5' RT primer (A-G-G-T-C-G-A-C-G-G-T-A-T-C-G-G, (SEQ ID NO: 64). In step two, 400 pg of cDNA product was used as PCR template in four separate reactions with each of the four 5' PCR primers of the form G-G-T-C-G-A-C-G-G-T-A-T-C-G-G-N (SEQ ID NO: 65), each paired with a "universal" 3' PCR primer G-A-G-C-T-C-C-A-C-C-G-C-G-G-T (SEQ ID NO: 66).

In step three, the product of each subpool was further divided into 64 subsubpools (2ng in 20µl) for the second PCR reaction, with 100 ng each of the fluoresceinated "universal" 3' PCR primer, the oligonucleotide (SEQ ID NO: 66) conjugated to 6-FAM and the appropriate 5' PCR primer of the form C-G-A-C-G-G-T-A-T-C-G-G-N-N-N-N (SEQ ID NO: 67), using a program that included an annealing step at a temperature X slightly above the T_m of each 5' PCR primer to minimize artifactual mispriming and promote high fidelity copying.

Each polymerase chain reaction step was performed in the presence of TaqStart antibody (Clonetech).

The products from the final polymerase chain reaction step for each of the tissue samples were resolved on a series of denaturing DNA sequencing gels using the automated ABI Prizm 377 sequencer. Data were collected using the GeneScan software package (ABI) and normalized for amplitude and migration. Complete execution of this series of reactions generated 64 product subpools for each of the four pools established by the 5' PCR primers of the first PCR reaction, for a total of 256 product subpools for the entire 5' PCR primer set of the second PCR reaction.

The mRNA samples from each timepoint after DSS treatment as described above were analyzed. Table 1 is a summary of the expression levels of 414 mRNAs determined from cDNA. These cDNA molecules are identified by their digital address, that is, a partial 5' terminus nucleotide sequence coupled with the length of the molecule, as well as the relative amount of the molecule produced at different time intervals after treatment. The 5' terminus partial nucleotide sequence is determined by the recognition site for MspI and the nucleotide sequence of the parsing bases of the 5' PCR primer used in the final PCR step. The digital address length of the fragment was determined by interpolation on a standard curve and, as such, may vary $\pm$ 1-2 b.p. from the actual length as determined by sequencing.

For example, the entry in Table 1 that describes a DNA molecule identified by the digital address MspI AGTG 244, is further characterized as having a 5' terminus partial nucleotide sequence of CGGAGTG and a digital address length of 244 b.p. The DNA molecule identified as MspI AGTG 244 is further described as being expressed at increasing levels at days 0, 4 and 8 with a moderate decline at day 12. However, the treatment results in a different pattern of expression of MspI AGTA 187, which declines on days 4 and 8 from the relatively high level seen at day 0, but increases at day 12.

Similarly, the other 412 DNA molecules identified in Table 1 by their MspI digital addresses are further characterized by the pattern of the level of gene expression on days 0, 4, 8 and 12 following the end of the DSS treatment. Many of the isolated clones were further

characterized in Tables 2 and 3. Their nucleotide sequences are provided as SEQ ID NO:1-62 in the Sequence Listing below.

The data shown in Figure 1 were generated with a 5'-PCR primer (C-G-A-C-G-G-T-A-T-C-G-G-C-G-C-G, SEQ ID NO: 68) paired with the "universal" 3' primer (SEQ ID NO: 66) labeled with 6-carboxyfluorescein (6FAM, ABI) at the 5' terminus. PCR reaction products were resolved by gel electrophoresis on 4.5% acrylamide gels and fluorescence data acquired on ABI377 automated sequencers. Data were analyzed using GeneScan software (Perkin-Elmer).

The results of TOGA analysis using a 5' PCR primer with parsing bases CGCG (SEQ ID NO: 68) are shown in Figure 1, which presents the results of TOGA analysis using a 5' PCR primer with parsing bases CGCG, showing PCR products produced from mRNA extracted from (top to bottom panels) colons isolated from mice 0 (Figure 1A), 4 (Figure 1B), 8 (Figure 1C) or 12 days (Figure 1D) after a seven day course of treatment with DSS. The vertical index line indicates a PCR product of about 458 b.p. that is present on day 0, reduced on day 4, much increased on day 8 and whose expression relatively decreases but is still elevated on day 12.

Some products, which were also differentially represented, appeared to migrate in positions that suggest that the products were novel based on comparison to data extracted from GenBank. In these cases, the PCR product was isolated, cloned into a TOPO vector (Invitrogen) and sequenced on both strands. Examples are found in Table 4, below. In order to verify that the clones isolated are from the same peak, PCR primers were designed based on the determined sequence and PCR was performed using the cDNA produced in the first PCR reaction as substrate. For example, for the 458 b.p. product disclosed above, oligonucleotides were synthesized using the universal 3' PCR primer and a 5' PCR primer corresponding to the 5' PCR primer in the second PCR step extended at the 3' end with additional nucleotides from the clone sequence 3' to the parsing bases (in this case, CGCG). This extended 5' PCR primer had the sequence: GATCGAATCC GGCGCGCACG GGGACCAGAC (SEQ ID NO: 78).

The products were separated by electrophoresis and the length of the clone was compared to the length of the original PCR product as shown in Figure 2. Figure 2A shows the length (as peak position) of the PCR product derived as described above. Figure 2B shows the PCR products produced using the original PCR primers, SEQ ID NO: 68 and SEQ ID NO: 66 (compare to Figure 1A). In Figure 2C, the traces from the top and middle panels are overlaid, demonstrating that the PCR product using the sequence of the isolated clone is the same length as the original PCR product.

The same procedure was used to verify candidate matches to database entries. The results are shown in Table 4, below. In each case, oligonucleotides were synthesized using the universal 3' PCR primer and a 5' PCR primer corresponding to the 5' PCR primer in the second PCR step extended at the 3' end with additional nucleotides from the sequences adjacent to the terminal MspI sites in the identified corresponding GenBank sequences. In three cases (IMX2_33, SEQ ID NO:21, IMX2_34, SEQ ID NO:61 and IMX2_70, SEQ ID NO:62) the DST sequence listed was obtained from the verified database match sequence in GenBank.

Figure 3 is a graphical representation of the results of Northern Blot analysis of clone IMX 2_46, SEQ ID NO: 10, where an agarose gel containing poly A enriched mRNA from the four experimental samples (0, 4, 8 or 12 days post-treatment) as well as size standards was blotted after electrophoresis and probed with radiolabelled IMX 2_46, SEQ ID NO: 10, imaged using a phosphorimager and quantified. Quantitative results showing the relative expression levels of the 1.6 kb transcript were: 0 day, 64; 4 days, 53; 8 days, 223; and 12 days, 269. The amount of RNA loaded on the gel was determined by probing for cyclophilin ("cyc").

TABLE 1							
Seq ID	Clone ID	Digital Address (Msp1)		0 Days	4 Days	8 Days	12 Days
		AAAA	157	405	247	181	459
		AAAA	316	42	34	47	174
51	IMX2_65	AAAC	186	3722	1457	4314	5810
50	IMX2_1	AAAC	407	95	106	345	281
		AAAG	449	214	121	249	299
		AACA	377	84	50	197	107
59	IMX2_2	AACG	299	79	108	366	91
		AATC	318	326	174	36	72
60	IMX2_3	AATG	114	210	136	1059	514
		ACAA	227	110	496	1568	564
		ACAA	381	51	57	120	69
		ACAC	283	131	69	212	143
1	IMX2_4	ACAC	320	28	38	284	106
26	IMX2_5	ACAC	398	29	37	94	70
		ACAG	288	69	67	264	101
		ACAT	335	767	362	231	449
		ACCA	359	652	134	227	452
		ACCC	493	128	332	316	398
		ACCG	387	218	60	149	570
		ACCT	225	851	189	650	1106
27	IMX2_6	ACCT	364	173	92	530	142
		ACGA	171	761	192	909	790
		ACGA	206	618	984	2071	1168
		ACGA	235	623	489	1511	1276
		ACGA	326	514	1217	805	722
		ACGC	361	455	97	227	299
		ACGG	198	1483	2094	5210	3294
52	IMX2_66	ACGG	238	1576	537	1476	1734
		ACGG	409	405	172	473	800
		ACGG	449	102	66	69	152
		ACGT	174	743	401	265	615
		ACTA	333	555	181	305	468
		ACTC	111	468	488	1773	1095
28	IMX2_7	ACTC	171	55	242	1326	651
		ACTG	184	1889	742	596	924
		ACTG	319	67	130	22	91
		ACTT	292	51	39	180	177
		AGAA	104	325	215	319	546
		AGAA	263	423	840	1050	1197
		AGAC	487	1568	826	354	1594
		AGAG	110	1692	2388	3206	1620
		AGAG	205	710	300	547	453

		AGAG	318	283	166	139	452
		AGAG	329	300	125	234	400
		AGAG	372	158	94	193	272
		AGCA	251	1125	353	659	1187
29	IMX2_8	AGCA	285	314	2300	1575	1050
		AGCA	316	158	656	179	174
		AGCC	379	439	148	366	544
		AGCG	258	278	176	543	239
		AGCG	284	163	39	25	87
		AGCG	396	476	57	261	605
		AGCT	175	920	1790	1538	1268
		AGCT	230	179	192	33	132
		AGGA	163	691	924	1919	830
		AGGA	190	1099	473	638	791
		AGGA	266	354	106	74	133
		AGGA	332	63	192	130	58
		AGGA	381	20	88	49	33
		AGGC	367	599	165	273	723
		AGGC	410	4628	1660	5080	4468
		AGGT	101	316	524	211	133
		AGGT	282	169	55	30	64
		AGGT	304	47	33	39	133
		AGGT	485	403	181	292	421
		AGTA	187	1331	412	457	753
		AGTA	209	194	331	365	412
		AGTA	285	62	257	269	164
		AGTC	288	101	354	197	163
		AGTC	324	665	319	66	112
30	IMX2_11	AGTC	476	51	150	661	373
31	IMX2_12	AGTG	244	165	338	2558	1077
		AGTT	101	92	235	234	70
		AGTT	248	70	164	303	124
		ATAA	170	410	116	241	315
41	IMX2_39	ATAA	226	38	44	174	69
		ATAA	261	485	282	891	766
32	IMX2_13	ATAC	226	174	303	598	221
		ATAC	275	954	756	1600	1133
		ATAG	110	4836	1250	1558	3142
		ATAG	151	1774	630	919	1402
		ATAG	223	513	584	1488	1078
		ATAG	323	1044	599	1180	1349
		ATAG	403	187	79	110	225
		ATCA	87	1018	1221	1588	664
		ATCA	121	671	263	160	314
		ATCA	466	376	150	251	298
		ATCC	224	89	175	382	232
		ATCC	304	489	245	553	537

		ATCC	347	229	106	283	291
		ATCC	451	309	112	187	295
		ATCG	137	399	111	147	135
		ATCG	205	67	240	341	76
		ATCG	362	61	44	165	94
		ATCG	464	393	139	330	394
		ATCT	392	208	72	123	236
		ATGA	203	743	1154	1968	1196
		ATGA	275	425	255	680	395
		ATGA	286	460	167	143	330
		ATGC	126	1310	551	609	711
		ATGG	196	284	470	636	361
		ATGG	349	481	588	178	439
		ATTA	105	432	233	67	94
		ATTA	162	568	201	235	399
		ATTA	263	104	48	51	69
		ATTG	140	1744	870	1490	2200
		CAAA	267	94	77	244	127
		CAAA	310	405	145	404	441
		CAAA	359	1777	461	594	1572
		CAAC	267	868	129	357	710
		CAAC	338	469	131	342	468
		CAAC	347	378	174	230	486
		CAAC	362	289	125	620	369
		CAAG	464	162	42	112	198
		CAAG	479	87	42	59	105
		CAAT	132	983	727	1005	1669
		CAAT	188	2436	1137	2640	3409
		CAAT	279	374	449	123	377
		CAAT	311	108	102	41	110
		CACA	392	150	55	41	137
		CACC	86	682	1597	1849	916
		CACC	360	85	92	280	89
		CACC	386	242	65	133	222
		CACC	394	45	47	135	39
		CACG	92	2188	4464	5216	2074
		CAGA	276	43	49	145	129
		CAGC	116	3026	6760	2728	3430
		CAGG	315	299	84	145	335
		CAGG	421	55	79	185	163
		CATC	162	1182	2890	1216	1262
		CATC	334	389	332	196	702
		CATC	368	176	57	103	162
		CATG	257	532	157	326	468
		CATG	353	2449	321	1050	1948
42	IMX2_40	CATT	90	299	2530	3750	823
		CATT	331	52	127	194	33

		CATT	347	449	72	261	393
		CCAA	126	2480	589	976	1713
		CCAA	166	1119	1451	960	545
43	IMX2_42	CCAA	201	242	1101	645	304
8	IMX2_43	CCAA	272	70	439	261	75
		CCAA	309	319	87	203	458
		CCAA	344	185	1389	992	201
		CCAC	178	85	295	222	122
		CCAC	248	1237	3609	891	1472
		CCAC	291	345	141	179	291
		CCAC	354	198	95	47	189
		CCAG	274	238	112	165	248
		CCAG	424	181	69	94	167
		CCAT	85	160	192	198	80
		CCAT	141	265	613	424	447
		CCAT	177	6480	4379	1011	1817
		CCCA	78	1043	1200	2301	1018
		CCCA	220	1438	773	454	463
		CCCA	255	29	27	23	79
		CCCG	102	2353	1931	3666	4085
		CCCG	211	1290	512	169	678
13	IMX2_55	CCCG	473	3711	1320	23	34
		CCCT	314	430	222	154	171
		CCCT	462	46	138	57	65
		CCGA	188	536	1180	948	647
		CCGA	207	297	803	357	340
		CCGC	114	1184	515	251	817
		CCGC	257	206	135	77	82
53	IMX2_68A	CCGC	266	286	77	289	435
54	IMX2_68B	CCGC	266	286	77	289	435
		CCGC	496	78	29	95	77
55	IMX2_69	CCGT	151	1603	264	440	1074
		CCGT	177	2085	836	411	338
		CCGT	214	1054	2000	260	512
		CCGT	252	682	1085	2385	2117
		CCGT	444	47	35	109	79
		CCGT	484	72	262	151	78
		CCTA	356	45	30	38	306
		CCTC	142	451	509	262	946
		CCTC	197	684	472	246	443
		CCTC	379	164	215	80	81
		CCTC	446	45	45	47	153
		CCTG	116	250	722	485	182
		CCTG	164	627	77	764	46
		CCTT	272	225	484	179	98
		CCTT	356	76	82	146	53
		CCTT	445	238	114	47	376

		CCTT	490	100	11	37	17
		CGAA	83	5360	4861	1634	4700
		CGAA	164	124	317	258	270
		CGAC	166	540	1253	1829	1039
		CGAC	267	503	79	680	609
		CGAC	364	169	65	415	336
		CGAG	262	23	208	146	600
		CGAG	302	373	49	217	335
		CGAG	352	120	49	155	185
15	IMX2_57	CGAT	176	1458	722	192	188
		CGCA	209	431	161	233	251
		CGCA	221	787	381	113	138
10	IMX2_46	CGCG	458	195	108	1280	793
		CGCG	460	230	89	1136	732
47	IMX2_58	CGCG	472	846	115	44	45
		CGCG	491	148	70	145	285
		CGGA	205	41	276	112	291
48	IMX2_59	CGGC	323	1873	1563	344	544
		CGGG	123	296	431	648	521
		CGGG	245	486	162	280	364
		CGGT	468	29	31	50	80
		CGTC	310	134	365	399	217
		CGTG	153	482	238	157	245
		CGTG	374	340	152	88	216
		CGTT	90	231	1068	1139	306
		CTAA	87	2094	817	719	2495
		CTAA	98	722	470	272	296
		CTAC	336	431	249	517	653
		CTAC	452	28	189	87	40
		CTAG	406	457	149	307	479
		CTAT	175	313	295	71	75
		CTAT	348	353	327	525	737
		CTCA	223	2614	726	2346	2193
		CTCA	242	66	74	172	146
		CTCA	329	143	63	84	150
		CTCA	347	179	62	57	178
		CTCA	391	3031	972	1564	2427
21	IMX2_33	CTCC	198	151	210	1699	2347
		CTCC	287	110	38	45	129
		CTCC	308	440	188	493	622
12	IMX2_48	CTCC	386	35	43	233	103
		CTCG	188	1648	225	1027	1373
		CTCG	428	157	56	220	262
49	IMX2_60	CTCG	472	486	28	15	20
		CTCT	117	111	370	213	161
		CTCT	407	59	37	107	91
		CTGC	147	300	763	888	484

		CTGC	257	908	711	2519	1665
		CTGC	414	463	70	223	550
		CTGG	200	1421	332	1230	1459
		CTGG	272	624	214	572	638
		CTGG	414	119	58	109	149
		CTTA	395	81	27	73	186
		CTTC	247	562	239	377	431
		CTTC	404	82	111	85	243
		CTTG	205	363	607	283	230
		CTTG	258	299	86	121	158
		CTTG	297	450	160	630	472
		CTTG	354	236	95	162	300
		CTTG	366	507	170	434	717
		CTTT	400	71	30	101	283
		CTTT	410	41	48	67	138
		CTTT	453	74	20	64	187
		GAAA	305	106	111	210	126
		GAAC	279	75	49	199	217
		GAAG	95	628	604	1402	449
		GACA	386	300	123	132	362
		GACA	400	802	295	412	1007
		GACC	119	217	771	217	202
		GACC	129	1056	3122	2325	1367
		GACG	89	1849	741	474	324
		GACG	169	502	538	464	179
		GACG	283	150	111	223	328
		GACG	381	43	44	161	53
		GACT	83	371	878	473	299
		GACT	315	894	554	823	1160
		GAGA	220	198	561	395	297
		GAGA	368	149	71	66	146
		GAGC	88	275	782	236	166
		GAGC	262	1703	474	511	1530
		GAGC	419	97	56	165	227
		GAGG	88	1550	2018	736	639
		GAGG	112	304	701	1046	788
		GAGG	253	143	59	75	107
		GAGT	197	1309	3415	1817	860
		GATA	339	200	131	32	42
		GATC	156	483	1229	764	551
		GATC	253	167	152	293	133
		GATC	450	434	180	389	477
24	IMX2_49	GATG	285	49	40	301	106
		GATT	126	3424	1173	2935	4890
		GATT	214	2246	692	1808	2796
		GCAA	226	173	755	823	572
		GCAA	333	177	41	116	152

		GCAG	118	337	434	654	1602
		GCAT	242	1111	1093	115	369
62	IMX2_70	GCAT	276	2425	1784	183	770
56	IMX2_71	GCAT	361	261	65	168	476
		GCAT	447	93	73	94	176
		GCCC	444	145	77	228	156
		GCCC	452	162	60	209	149
		GCCG	269	770	28	359	512
33	IMX2_15	GCCG	364	72	120	378	206
34	IMX2_16	GCGA	190	169	539	392	140
		GCTA	82	1049	3530	2007	819
		GCTA	269	1387	490	598	1485
		GCTA	452	41	261	150	42
		GCTC	157	1515	467	659	1151
3	IMX2_17	GCTC	245	258	864	1457	514
57	IMX2_72	GCTC	425	460	67	208	650
		GCTG	160	1171	2011	1804	726
		GGAA	416	1601	1067	553	2083
		GGAC	270	62	231	162	93
		GGAC	327	527	145	409	727
		GGAG	274	309	83	398	472
		GGAG	283	266	67	215	306
		GGAG	440	73	38	73	98
		GGAT	276	1016	639	118	259
		GGAT	362	140	35	55	152
		GGCA	341	155	94	56	168
		GGCA	349	214	63	88	192
		GGCC	170	920	103	257	467
		GGCC	327	86	345	470	87
		GGCT	445	447	79	403	606
		GGGA	430	817	139	777	1212
		GGGC	142	1375	565	529	389
		GGGC	355	288	61	205	271
		GGGC	418	53	44	64	99
		GGGG	241	126	389	812	371
35	IMX2_20	GGGT	150	138	190	461	328
61	IMX2_34	GGGT	177	622	1213	3329	2197
		GGTC	117	1171	2648	2360	1133
		GGTC	374	104	72	295	285
		GGTT	268	73	41	22	129
		GTAG	117	204	331	299	1780
		GTAG	437	57	66	146	118
		GTAT	242	407	356	18	73
17	IMX2_61	GTAT	276	896	608	39	189
		GTAT	362	184	67	106	257
		GTAT	382	132	25	30	121
		GTAT	409	40	32	27	116

		GTCA	92	555	813	1400	722
		GTGA	88	968	2297	1809	1928
		GTGC	293	565	160	147	541
		GTGC	442	730	204	705	1245
		GTGT	240	31	24	22	160
		GTGT	276	65	59	97	143
		GTTC	167	104	172	323	318
		GTTC	490	35	24	72	33
		GTTT	198	70	86	203	64
		TAAC	92	144	162	500	167
		TAAC	331	122	45	68	63
		TAAC	395	47	52	104	117
		TAAG	143	88	120	335	162
23	IMX2_21	TAAG	192	129	46	92	594
		TAAG	201	59	120	193	150
		TAAG	244	278	118	107	133
		TAAG	266	44	26	37	90
58	IMX2_73	TAAG	302	4180	1203	490	1807
		TAAT	267	59	131	88	178
		TACC	343	54	91	242	49
		TACC	415	206	218	446	439
		TACG	284	1286	411	984	1079
		TACT	160	1169	944	656	533
40	IMX2_35	TACT	338	235	316	850	781
		TAGA	121	329	139	52	157
		TAGC	219	1774	2833	577	4624
44	IMX2_51	TAGG	80	392	2299	2373	469
		TAGG	175	2013	846	314	489
		TAGT	177	858	886	199	141
4	IMX2_22	TAGT	236	261	118	1534	1085
		TATC	314	379	292	115	241
		TATT	268	80	153	112	238
		TCAG	223	3378	931	1638	3793
36	IMX2_23	TCAT	161	314	800	1426	849
		TCAT	176	1442	802	372	371
6	IMX2_36	TCAT	188	99	856	1141	921
		TCCA	312	430	51	267	553
		TCCC	314	468	74	316	519
		TCCG	472	388	54	24	43
		TCCT	176	1100	1002	2127	1547
25	IMX2_62	TCCT	259	618	175	281	347
20	IMX2_74	TCTC	232	1700	164	230	653
37	IMX2_24	TCTC	365	94	118	819	435
		TCTC	392	27	62	199	126
		TCTG	120	907	1918	925	586
		TCTG	154	219	437	813	345
		TCTG	263	138	301	612	292

		TCTG	381	113	15	17	28
		TCTT	197	396	1276	779	660
		TGAC	175	365	45	54	36
		TGAC	283	340	175	157	71
38	IMX2_25	TGAC	376	9	38	394	448
		TGAC	416	2068	581	787	2274
		TGAT	279	193	129	286	664
19	IMX2_63	TGAT	476	2230	1330	42	139
		TGCA	79	1129	2391	2202	689
		TGCA	124	1783	1512	1373	982
		TGCA	186	203	147	1013	628
		TGCA	217	471	163	66	44
		TGCA	414	35	29	52	120
		TGCC	301	312	267	35	135
39	IMX2_26	TGCC	344	372	598	2223	563
		TGCG	120	186	215	683	581
		TGCG	410	2776	258	1208	2951
		TGCT	89	1267	2754	2523	928
		TGCT	403	61	105	280	208
		TGGA	95	1917	1923	1911	839
		TGGC	80	755	1969	2468	1069
		TGGC	121	185	1280	1415	455
22	IMX2_64	TGGC	175	1747	262	427	73
		TGGC	219	2529	5287	5480	5240
		TGGC	402	185	36	128	104
		TGGT	175	946	288	215	213
5	IMX2_28	TGTA	166	194	554	102	157
		TGTC	154	1177	918	3439	3334
		TGTC	264	5701	4480	1159	2280
45	IMX2_52	TGTC	343	38	26	296	32
		TGTG	266	67	664	294	187
		TGTG	362	1800	943	370	1050
		TTAA	313	140	288	281	340
46	IMX2_53	TTCC	320	29	515	354	109
		TTCG	346	37	255	100	32
		TTCT	267	59	88	98	146
		TTGA	161	161	239	803	467
		TTGC	353	54	92	220	109
		TTGC	383	37	201	573	255
		TTGG	289	139	116	541	324
		TTGT	264	144	264	465	325
		TTTC	149	287	1262	979	147
		TTTT	268	93	150	130	256

TABLE 2

Seq ID	Clone ID	Digital Address (MspI)	Database Match (Accession #)	Relative DST Amount*				Validation Method
				0 Days	4 Days	8 Days	12 Days	
59	IMX2_2	AACG 299	Mouse complement component C3 mRNA (K02782)	382	725	711	705	R
60	IMX2_3	AATG 114	Mouse mRNA for ribosomal protein L7 (X57960)	176	191	235	332	R
1	IMX2_4	ACAC 320	EST Soares mouse 3NbMS cDNA (AA178128)	669	1306	1281	810	N
28	IMX2_7	ACTC 171	Mouse monokine induced by gamma interferon (M34815)	79	87	722	796	R
30	IMX2_11	AGTC 476	Mouse complement component C3 mRNA (K02782)	207	650	671	517	R
33	IMX2_15	GCCG 364	Mouse FVB/N collagen pro-alpha (U08020)	370	667	565	717	R
23	IMX2_21	TAAG 192	NOVEL	240	408	533	866	N
4	IMX2_22	TAGT 236	Mouse mRNA for serine/theronine kinase (Y09010)	2931	3158	3749	4025	N
37	IMX2_24	TCTC 365	Mouse EN-7 mRNA (X53247)	117	136	155	291	N
38	IMX2_25	TGAC 376	Mouse mRNA for Ig heavy chain (X03690)	136	103	185	548	R
5	IMX2_28	TGTA 166	NOVEL	993	1195	727	699	N
21	IMX2_33	CTCC 198	Mouse secretory leukocyte protease inhibitor mRNA (U73004)	82	331	291	386	R
6	IMX2_36	TCAT 188	EST Barstead mouse proximal colon MPLRB6 Mus musculus cDNA (AA529850)	106	148	247	465	N
42	IMX2_40	CATT 90	EST Mus musculus unfertilized egg (AU023455)	51	216	305	56	R
8	IMX2_43	CCAA 272	EST Knowles Solter mouse blastocyst B1 Mus musculus cDNA clone (AA571661)	59	114	95	61	N
10	IMX2_46	CGCG 458	EST Soares mouse lymph node NbMLN Mus musculus cDNA clone 750847 (AA290194)	64	53	223	269	N
12	IMX2_48	CTCC 386	Mouse mRNA for macrophage inflammatory protein -1 (MIP2) (X53798)	See Figure 4				R

EST = Expressed Sequence Tag, NA = Not Applicable, R = Reverse Transcriptase-Polymerase Chain Reaction, N = Northern Blot Analysis *Relative DST Amount = Relative DST Amount for Northern blots have been corrected for background and normalized to the hybridization signal of cyclophilin. The values were obtained by exposing the Northern blot to a phosphorimaging screen and quantitated using the Phosphorimager SI.

TABLE 2 (continued)

Seq ID	Clone ID	Digital Address (Msp1)	Database Match (Accession #)	Relative DST Amount*				Validation Method
				0 Days	4 Days	8 Days	12 Days	
13	IMX2_55	CCCG 473	EST Barstead mouse irradiated colon MPLRB7 (AA823573)	194	45	0	7	R
49	IMX2_60	CTCG 472	EST Barstead mouse irradiated colon MPLRB7 (AA823573)	549	62	11	21	R
17	IMX2_61	GTAT 276	EST Barstead mouse pooled organs (AA240177)	297	82	401	524	R
22	IMX2_64	TGGC 175	NOVEL	238	0	0	0	N
57	IMX2_72	GCTC 425	Homolog to rat G protein gamma-5 Subunit mRNA (M95780)	191	226	155	173	N
20	IMX2_74	TCTC 232	EST Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA clone (AA689792)	See Figure 5				R

TABLE 3

Seq ID	Clone ID	Digital Address (MspI)	Database Match (Accession #)	% Homology	Nucleotide homology	
					DST nucleotide range (bp#)	Database nucleotide range (bp#)
50	IMX2_1	AAAC 407	Mouse ion channel homolog RIC mRNA (U72680)	99%	1 – 314	475 – 788
59	IMX2_2	AACG 299	Mouse complement component C3 mRNA (K02782)	99%	1 – 248	3456 – 3703
60	IMX2_3	AATG 114	Mouse mRNA for ribosomal protein L7 (X57960)	96%	1 – 64	545 – 608
1	IMX2_4	ACAC 320	EST Soares mouse 3NbMS cDNA (AA178128)	99%	1 – 264	178 – 441
26	IMX2_5	ACAC 398	Mouse MHC class II protein alpha chain, complete cds (M95514)	98%	1 – 348	653 – 999
27	IMX2_6	ACCT 364	Murine AIDS virus-related provirus (S80082)	99%	1 – 310	2716 – 3025
28	IMX2_7	ACTC 171	Mouse monokine induced by gamma interferon (M34815)	83%	1 – 117	321 – 437
29	IMX2_8	AGCA 285	Mus musculus 45S pre rRNA gene (X82564)	98%	1 – 234	6078 – 6309
30	IMX2_11	AGTC 476	Mouse complement component C3 mRNA (K02782)	97%	1 – 414	4673 – 5087
31	IMX2_12	AGTG 244	Mouse mRNA for serum amyloid (X03479)	100%	1 – 185	236 – 420
32	IMX2_13	ATAC 226	Mus musculus l-kappa B alpha chain mRNA (U36277)	100%	1 – 173	855 – 1027
33	IMX2_15	GCCG 364	Mouse FVB/N collagen pro-alpha (U08020)	99%	1 – 308	3616 – 3923
34	IMX2_16	GCGA 190	EST Mus musculus 8-cell embryo cDNA 3'end sequence (AU020205)	93%	1 – 124	11 – 134
3	IMX2_17	GCTC 245	Mouse CRP-ductin-alpha mRNA (U37438)	98%	1 – 192	2815 – 3006
35	IMX2_20	GGGT 150	Mouse MHC class III H2-C4 complement component (M12969)	98%	1 – 99	31 – 129
23	IMX2_21	TAAG 192	NOVEL	N/A	N/A	N/A
4	IMX2_22	TAGT 236	Mouse mRNA for serine/theronine kinase (Y09010)	100%	1 – 169	2555 – 2723

EST = Expressed Sequence Tag, N/A = Not Applicable

TABLE 3 (continued)

Seq ID	Clone ID	Digital Address (MspI)	Database Match (Accession #)	% Homology	Nucleotide homology	
					DST nucleotide range (bp#)	Database nucleotide range (bp#)
36	IMX2_23	TCAT 161	Mouse glycoprotein CD28 mRNA MUSCD28 (M34563)	99%	1 – 109	1382 – 1490
37	IMX2_24	TCTC 365	Mouse EN-7 mRNA (X53247)	98%	1 – 306	919 – 1222
38	IMX2_25	TGAC 376	Mouse mRNA for Ig heavy chain (X03690)	99%	1 – 324	494 – 817
39	IMX2_26	TGCC 344	Mouse DNA for alpha globin (X05379)	99%	1 – 291	644 – 1068
5	IMX2_28	TGTA 166	NOVEL	N/A	N/A	N/A
40	IMX2_35	TACT 338	Stratagene mouse macrophage (#937306) Mus musculus cDNA clone (AA919822)	98%	1 – 281	61 – 341
6	IMX2_36	TCAT 188	EST Barstead mouse proximal colon MPLRB6 Mus musculus cDNA (AA529850)	99%	1 – 133	336 – 467
41	IMX2_39	ATAA 226	Mouse ceruloplasmin mRNA (U49430)	99%	1 – 172	558 – 729
42	IMX2_40	CATT 90	EST Mus Musculus unfertilized egg (AU023455)	100%	1 – 36	117 – 152
43	IMX2_42	CCAA 201	EST mouse 4-cell embryo (AU041854)	96%	1 – 108	1 – 108
8	IMX2_43	CCAA 272	EST Knowles Solter mouse blastocyst B1 Mus musculus cDNA clone (AA571661)	98%	1 – 221	282 – 502
10	IMX2_46	CGCG 458	EST Soares mouse lymph node NbMLN Mus musculus cDNA clone 750847 (AA290194)	99%	1 – 298	175 – 471
12	IMX2_48	CTCC 386	Mouse mRNA for macrophage inflammatory protein –1 (MIP2) (X53798)	99%	1 – 331	65 – 395
24	IMX2_49	GATG 285	NOVEL	N/A	N/A	N/A
44	IMX2_51	TAGG 80	Stratagene mouse skin (#937313) Mus musculus cDNA clone (AI644131)	100%	1 – 24	23 – 46
45	IMX2_52	TGTC 343	Mouse DNA for alpha globin (X05379)	99%	1 – 291	644 – 1068
46	IMX2_53	TTCC 320	Mus musculus short incubation prion protein PrnPa gene, complete cds (U29186)	99%	1 – 283	21054 – 21336
13	IMX2_55	CCCG 473	EST Barstead mouse irradiated colon MPLRB7 (AA823573)	97%	5 – 414	37 – 445

TABLE 3 (continued)

Seq ID	Clone ID	Digital Address (MspI)	Database Match (Accession #)	% Homology	Nucleotide homology	
					DST nucleotide range (bp#)	Database nucleotide range (bp#)
15	IMX2_57	CGAT 176	EST Soares mouse NML cDNA (AA244542)	96%	1 -- 125	125 – 249
47	IMX2_58	CGCG 472	EST Barstead mouse irradiated colon MPLRB7 (AA823573)	96%	8 – 226	40 – 258
48	IMX2_59	CGGC 323	EST Soares mouse NML cDNA (AA268434)	99%	13 – 256	65 – 308
49	IMX2_60	CTCG 472	EST Barstead mouse irradiated colon MPLRB7 (AA823573)	97%	224 – 415	255 – 446
17	IMX2_61	GTAT 276	EST Barstead mouse pooled organs (AA240177)	99%	1 – 226	277 – 502
25	IMX2_62	TCCT 259	NOVEL	N/A	N/A	N/A
19	IMX2_63	TGAT 476	EST UI-M-BH3-avd-b-10-0-UI.s1 NIH_BMAP_M_S4 Mus musculus cDNA clone (AW495679)	99%	1 – 427	12 – 437
22	IMX2_64	TGGC 175	NOVEL	N/A	N/A	N/A
51	IMX2_65	AAAC 186	Mouse mitochondrial genes coding for tRNAs (V00665)	98%	4 – 141	1451 – 1588
52	IMX2_66	ACGG 238	Human RNA for RNA polymerase (Z23102)	91%	79 – 145	1428 – 1494
53	IMX2_68A	CCGC 266	Mouse mRNA regulated by bone morphogenic protein (X95281)	96%	1 – 207	652 – 858
54	IMX2_68B	CCGC 266	Mouse cytochrome beta-558 (M31775)	98%	1 – 196	468 – 664
55	IMX2_69	CCGT 151	Mouse beta-galactoside-Binding lectin (M33215)	95%	1 – 97	524 – 620
56	IMX2_71	GCAT 361	EST Soares mouse 3NME12 cDNA (AA199089)	99%	1 – 307	45 – 350
57	IMX2_72	GCTC 425	Homolog to rat G protein gamma-5 Subunit mRNA (M95780)	95%	1 – 221	88 – 308
58	IMX2_73	TAAG 302	EST Soares 2NbMT cDNA clone (AA116528)	100%	1 – 244	281 – 524
20	IMX2_74	TCTC 232	EST Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA clone (AA689792)	95%	1 – 171	175 – 341

TABLE 4: VERIFIED CANDIDATE MATCHES

Seq ID	Clone ID	Digital Address (Msp1)	Gene Identity (Accession #)	Extended Primer (SEQ ID NO:)
50	IMX2_1	AAAC 407	Mouse ion channel homolog RIC mRNA (U72680)	GAT CGA ATC CGG AAA CGG GGA CCG CTG GTG (SEQ ID NO:124)
26	IMX2_5	ACAC 398	Mouse MHC class II protein alpha chain, complete cds (M95514)	GAT CGA ATC CGG ACA CCA TAG AGA CCC TGA (SEQ ID NO:85)
27	IMX2_6	ACCT 364	Murine AIDS virus-related provirus (S80082)	GAT CGA ATC CGG ACC TCA CCG ACC AGC CCA (SEQ ID NO:125)
29	IMX2_8	AGCA 285	Mus musculus 45S pre rRNA gene (X82564)	GAT CGA ATC CGG AGC ACC ACA TCG ATC TAA (SEQ ID NO:86)
31	IMX2_12	AGTG 244	Mouse mRNA for serum amyloid (X03479)	GAT CGA ATC CGG AGT GGC AAA GAC CCC AAC (SEQ ID NO:126)
32	IMX2_13	ATAC 226	Mus musculus I-kappa B alpha chain mRNA (U36277)	GAT CGA ATC CGG ATA CAG CAG CAG CTG GGC (SEQ ID NO:69)
34	IMX2_16	GCGA 190	EST Mus musculus 8-cell embryo cDNA 3'end sequence (AU020205)	GAT CGA ATC CGG GCG ATG GTG GTG TAT GCC (SEQ ID NO:87)
3	IMX2_17	GCTC 245	Mouse CRP-ductin-alpha mRNA (U37438)	GAT CGA ATC CGG GCT CTG GGT CTA TTG TTC (SEQ ID NO:70)
35	IMX2_20	GGGT 150	Mouse MHC class III H2-C4 complement component (M12969)	GAT CGA ATC CGG GGG TGC CAG GTG TGA GGC (SEQ ID NO:71)
36	IMX2_23	TCAT 161	Mouse glycoprotein CD28 mRNA MUSCD28 (M34563)	GAT CGA ATC CGG TCA TGG GAA CTC AGT ATT (SEQ ID NO:72)
39	IMX2_26	TGCC 344	Mouse DNA for alpha globin (X05379)	GAT CGA ATC CGG TGC CCT GTC TGC TCT GAG (SEQ ID NO:73)
21	IMX2_33	CTCC 198	Mouse secretory leukocyte protease inhibitor mRNA (U73004)	GAT CGA ATC CGG CTC CCT GTA TCC CAG GCT (SEQ ID NO:74)
61	IMX2_34	GGGT 177	Mouse MHC sex-linked protein (M21576)	GAT CGA ATC CGG GGG TGC CAG GTG TGAGGC (SEQ ID NO:75)
40	IMX2_35	TACT 338	Stratagene mouse macrophage (#937306) Mus musculus cDNA clone (AA919822)	GAT CGA ATC CGG TAC TGG GGA GGC ACA GGC (SEQ ID NO:88)
41	IMX2_39	ATAA 226	Mouse ceruloplasmin mRNA (U49430)	GAT CGA ATC CGG ATA ACA GTA TGT GTA TGT (SEQ ID NO:76)
43	IMX2_42	CCAA 201	EST mouse 4-cell embryo (AU041854)	GAT CGA ATC CGG CCA AAC TCT CAA TTA CCA (SEQ ID NO:77)

EST = Expressed Sequence Tag

TABLE 4 (continued): VERIFIED CANDIDATE MATCHES

Seq ID	Clone ID	Digital Address (Msp1)	Gene Identity (Accession #)	Extended Primer (SEQ ID NO:)
10	IMX2_46	CGCG 458	EST Soares mouse lymph node NbMLN Mus musculus cDNA clone 750847 (AA290194)	GAT CGA ATC CGG CGC GCA CGG GGA CCA GAC (SEQ ID NO:78)
24	IMX2_49	GATG 285	NOVEL	GAT CGA ATC CGG GAT GTG GGA AGG TTA GAA (SEQ ID NO:89)
44	IMX2_51	TAGG 80	Stratagene mouse skin (#937313) Mus musculus cDNA clone (AI644131)	GAT CGA ATC CGG TAG GGT AGA GTG TCG CCA (SEQ ID NO:90)
45	IMX2_52	TGTC 343	Mouse DNA for alpha globin (X05379)	GAT CGA ATC CGG TGT CCT GTC TGC TCT GAG (SEQ ID NO:79)
46	IMX2_53	TTCC 320	Mus musculus short incubation prion Protein PrnPa gene, complete cds (U29186)	GAT CGA ATC CGG TTC CCA TAT CTT TGA GGG (SEQ ID NO:91)
15	IMX2_57	CGAT 176	EST Soares mouse NML cDNA (AA244542)	GAT CGA ATC CGG CGA TGT ACA CTC GGG TCA (SEQ ID NO:92)
47	IMX2_58	CGCG 472	EST Barstead mouse irradiated colon MPLRB7 (AA823573)	GAT CGA ATC CGG CGC GTA TCT GTG TGA ACT (SEQ ID NO:93)
48	IMX2_59	CGGC 323	EST Soares mouse NML cDNA (AA268434)	GAT CGA ATC CGG CGG CGA TAT CCA GTC TGG (SEQ ID NO:94)
25	IMX2_62	TCCT 259	NOVEL	GAT CGA ATC CGG TCC TGG CAG ACA GAC ATG (SEQ ID NO:95)
19	IMX2_63	TGAT 476	EST UI-M-BH3-avd-b-10-0-UI.s1 NIH_BMAP_M_S4 Mus musculus cDNA clone (AW495679)	GAT CGA ATC CGG TGA TAA GAG CAA CTT CGC (SEQ ID NO:96)
51	IMX2_65	AAAC 186	Mouse mitochondrial genes Coding for tRNAs (V00665)	GAT CGA ATC CGG AAA CCC CGA AAC CAA ACG (SEQ ID NO:80)
52	IMX2_66	ACGG 238	Human RNA for RNA polymerase (Z23102)	GAT CGA ATC CGG ACG GAG GAC CAC CCG TGC (SEQ ID NO:81)
53	IMX2_68A	CCGC 266	Mouse mRNA regulated by bone morphogenic protein (X95281)	GAT CGA ATC CGG CCG CCA CCC AAC AAC TTT (SEQ ID NO:97)
54	IMX2_68B	CCGC 266	Mouse cytochrome beta-558 (M31775)	GAT CGA ATC CGG CCG CCC GCA GAG GTC CGA (SEQ ID NO:98)
55	IMX2_69	CCGT 151	Mouse beta-galactoside-Binding Lectin (M33215)	GAT CGA ATC CGG CCG TGT GTG CCG TAG GAG (SEQ ID NO:82)
62	IMX2_70	GCAT 276	Mouse mRNA for preproelastase II (X04573)	GAT CGA ATC CGG GCA TCT AAT GGC CAG TGG (SEQ ID NO:83)
56	IMX2_71	GCAT 361	EST Soares mouse 3NME12 cDNA (AA199089)	GAT CGA ATC CGG GCA TCC ATG GGT TCC AAC (SEQ ID NO:84)
58	IMX2_73	TAAG 302	EST Soares 2NbMT cDNA clone (AA116528)	GAT CGA ATC CGG TAA GCA TGG CAA GAC CCG (SEQ ID NO:99)

EST = Expressed Sequence Tag

TABLE 5: RT-PCR Validation Primers

Seq ID	Clone ID	Digital Address (Msp1)	Database Match (Accession #)	5' RT-PCR Primer and 3' RT-PCR Primer
59	IMX2_2	AACG 299	Mouse complement component C3 mRNA (K02782)	CACAGCCTTC GTCCTCATCG CACTG (SEQ ID NO:100) TTG TTCATCA GGGCCAGGGC ATACC (SEQ ID NO:101)
60	IMX2_3	AATG 114	Mouse mRNA for ribosomal protein L7 (X57960)	TCTGAAGCCC CGTGCTCCAC CCACT (SEQ ID NO:102) TCACGGCCCC GCTCCCATTC C (SEQ ID NO:103)
28	IMX2_7	ACTC 171	Mouse monokine induced by gamma interferon (M34815)	CCAAGTCCCA GGCCTGTCTG TT (SEQ ID NO:104) TGGTCTCCAC TGTAGAACC CCAAAA (SEQ ID NO:105)
30	IMX2_11	AGTC 476	Mouse complement component C3 mRNA (K02782)	ACATAGAGCT GTTGGATGAT TTTGA (SEQ ID NO:106) CAAGTTCTTC GCACTGTTTC TGGTA (SEQ ID NO:107)
33	IMX2_15	GCCG 364	Mouse FVB/N collagen pro-alpha (U08020)	CGACCTCAAG ATGTGCCACT CTGA (SEQ ID NO:108) ACCAGTTCTT CTGAGGCACA GACGG (SEQ ID NO:109)
38	IMX2_25	TGAC 376	Mouse mRNA for Ig heavy chain (X03690)	GAACAAAGGA TCCACACCCC AAACC (SEQ ID NO:110) GCACATGTGG AGGACACGTT CTTCA (SEQ ID NO:111)
21	IMX2_33	CTCC 198	Mouse secretory leukocyte protease inhibitor mRNA (U73004)	CAGTGTGGAG GAAGCCTGGG AGGTG (SEQ ID NO:127) CACATCGGGG GCAGGCAGAC TTTC (SEQ ID NO:128)
42	IMX2_40	CATT 90	EST Mus musculus unfertilized egg (AU023455)	ATGAAAAATA TGGAAAATGA TAAAA (SEQ ID NO:112) CTAAAATGTT CTACAGTGTG GTTT (SEQ ID NO:113)
12	IMX2_48	CTCC 386	Mouse mRNA for macrophage inflammatory protein -1 (MIP2) (X53798)	GCCCAGACAG AAGTCATAGC CACTC (SEQ ID NO:114) TTTATGGTTC TTCCGTTGAG GGACA (SEQ ID NO:115)
13	IMX2_55	CCCG 473	EST Barstead mouse irradiated colon MPLRB7 (AA823573)	GAGTCTGGCT CTGGGATTGC AGAA (SEQ ID NO:116) CCCCCATAGG AATCCTGCTC TTGT (SEQ ID NO:117)
49	IMX2_60	CTCG 472	EST Barstead mouse irradiated colon MPLRB7 (AA823573)	CCACTGGGGA GGAAAGGCTT GAA (SEQ ID NO:118) CCACATGGTC CCACAAGGAA ACATC (SEQ ID NO:119)
17	IMX2_61	GTAT 276	EST Barstead mouse pooled organs (AA240177)	GCAGGTGCAT GGCATCGTGA (SEQ ID NO:120) GGGGACAGTG GCAGTAATGT CTTCA (SEQ ID NO:121)
20	IMX2_74	TCTC 232	EST Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA clone (AA689792)	TCAGAGATTA GCATGGTGGG ACA (SEQ ID NO:122) CTGGTTTGAC AGAGACGCAG TAGTC (SEQ ID NO:123)

EXAMPLE 2

RT-PCR Validation

RT-PCR validation of cloned DSTs was performed, and results are presented in Table 2. The starting amount of template was chosen based on a control curve that was generated to accurately define the linear range of amplication for the given cloned DST. Based upon the intensity of the peak cloned from the TOGA™ panel, the following amounts of template were chosen:

TOGA peak intensity	Low [cDNA]	High [cDNA]
0 – 400	2000 pg	5000 pg
300-1000	400 pg	2000 pg
> 1000	50 pg	250 pg

The PCR primers used for validation of cloned DSTs are listed in Table 5. Duplicate reaction mixtures were assembled using the appropriate low and high concentration of cDNA template chosen from the time point sample showing the strongest TOGA™ signal. The reaction mixtures were cycled for 23, 26, 29, 32, 35, and 38 cycles. The resulting amplification products from the duplicate reactions were quantitated and plotted against the cycle number to generate a standard curve. From these data, the cycle number and cDNA concentration combination which yielded acceptable levels of PCR product within the linear range of amplification were chosen for RT-PCR validation across the various time-points.

The RT-PCR validation consisted of assembling triplicate reactions using the chosen concentration of cDNA cycled to the defined cycle number. For example, the data in Table 2 for IMX2_55 (SEQ ID NO:13) were generated using 50 pg cDNA template and 29 cycles. An internal control primer pair amplified under the same conditions was also performed to provide the basis for normalizing any differences between the cDNA templates.

EXAMPLE 3

RT-PCR Analysis Using Fluorimetry

Two DSTs were validated using an alternative protocol. The primers used for RT-PCR are listed in Table 5. For each DST examined, the optimal annealing temperature and reagent conditions were determined for the corresponding set of primers (see Table 5) based on results from a preliminary experiment. In eight separate reactions, each set of primers was assayed to find the optimal conditions by adjusting the following four parameters: primer concentration, dNTP concentration, $MgCl_2$ concentration, and Taq polymerase. Once optimal conditions were determined, each DST was run in duplicate multiple simultaneous reactions which usually included at least four dilutions of template, plus control reactions lacking template, and six sequential data points for numbers of cycles.

Reactions were performed using "Hot Start" PCR with the Clontech TaqStart antibody system (Cat. #5400-1). Each reaction contained 1µl of the cDNA library dilution as template, determined amounts of AmpliTaq DNA polymerase (cat. #N808-0156), $MgCl_2$, dNTPs (GibcoBRL cat. #10297-018), primer, and Clontech TaqStart Antibody in a 20µl reaction volume using 10x Taq buffer II (without $MgCl_2$). Typically, a master mix containing all components except the template was prepared and aliquoted. Various templates were then added to these master mix samples and 20 µl volumes were subsequently dispensed into individual reaction tubes. During the PCR run, tubes were removed sequentially on a predetermined schedule in order to quantitate expression of the target DST over a "window" of cycles. After amplification, the samples were quantified via fluorimetry.

PCR was performed at annealing temperatures about 5 degrees above the lowest melting temperature of each primer pair using the following program: 1) 95 degrees Celsius, 3 minutes; 2) 95 degrees Celsius, 30 seconds; 3) TM+5 degrees Celsius, 30 seconds; 4) 72 degrees Celsius, for a time dependent on target length at 16 bp/second; 5) repeat steps 2-4 33 more cycles; 6) 72 degrees Celsius, 3 minutes; 7) 14 degrees Celsius, forever.

Following temperature cycling, 2µl of the PCR reaction was added to 140µl of a 1:280 dilution of PicoGreen (Molecular Probes cat. #P-11495 (10x100µl)) in TE pH 7.5 in a 96-well Costar UV microtiter plate (Fisher cat. #07-200623). The samples were mixed gently for 1.5 minutes and allowed to equilibrate at room temperature in the dark for 15 minutes. The concentration of the PCR products was quantified by fluorimetry using a PerSeptive Biosystems CytoFluor series 4000 multi-well plate reader.

Background fluorescence was determined by using duplicate control samples that were cycled with all reaction components except the template. The mean value from these duplicate background control samples was subtracted from the corresponding experimental values prior to analyzing results. The sensitivity of the PicoGreen dsDNA assay is reported to be 250pg/ml (50pg dsDNA in a 200µl assay volume) using a fluorescence microplate reader such as was used in these measurements.

The results of the quantified RT-PCR of IMX 2_48 for a 1:2000 dilution of the library are shown in Figure 4 (in arbitrary fluorescence units) and in Table 6 (normalized to the control value at each time point). The results of the quantified RT-PCR of IMX 2_74 for a 1:500 dilution of the library are shown in Figure 5 (in arbitrary fluorescence units) and in Table 6 (normalized to the control value at each time point).

Table 6

IMX2_48						IMX2_74				
Cycle						cycle				
	1:2000 0d	1:2000 4d	1:2000 8d	1:2000 12d			1:500 0d	1:500 4d	1:500 8d	1:500 12d
24	1.0	0.9	1.0	1.0						
26	1.0	2.4	2.0	1.4		26	1.00	0.17	0.12	0.38
28	1.0	3.3	3.4	3.4						
						29	1.00	0.25	0.33	0.44
30	1.0	4.5	6.7	6.1						
32	1.0	3.9	10.2	11.3		32	1.00	0.60	0.56	0.70
34	1.0	15.5	36.0	36.5						
TOGA	1.0	1.6	7.3	2.8		TOGA	1.00	0.10	0.14	0.38

Table 7

Relationships of Extended Sequences

Seq ID	Extended Sequence ID	DST alignment to Extended Sequence (bp #)	Database Match (Accession #)	% Homology	Nucleotide homology	
					Nucleotide range of extended sequence (bp #)	Database nucleotide range (bp #)
2	IMX2_4.EXT	688 – 957	Homology to Homo sapiens chromosome 19, cosmid R27656, complete sequence (AC007565.1)	76% 71% 73% 83% 79% 62% 59% 75% 67% 69% 62% 75% 77%	2 – 97 53 – 200 214 – 282 255 – 297 297 – 450 529 – 646 714 – 887 792 – 852 792 – 871 792 – 891 796 – 916 884 – 935 884 – 935	7692 – 7785 15459 – 15618 19159 – 19225 22411 – 22453 28497 – 28651 28700 – 28820 8337 – 8515 41775 – 41835 16338 – 16414 39753 – 39848 36657 – 36785 18855 – 18907 25677 – 25729
7	IMX2_36.EXT	293 – 427	Mus musculus DAP12 mRNA, complete cds (AF024637.1) Mus musculus KAR-associated protein mRNA, complete cds (AF077829.1)	100% 100%	1 – 254 1 – 254	92 – 345 92 – 345
9	IMX2_43.EXT	1097 – 1298	Homology not determined			
11	IMX2_46.EXT	157 – 561	Homology to Homo sapiens regulator of Fas-induced apoptosis (TOSO) mRNA (NM_005449.1)	74%	113 – 370	913 – 1170

Table 7, continued

Seq ID	Extended Sequence ID	DST alignment to Extended Sequence (bp #)	Database Match (Accession #)	% Homology	Nucleotide homology	
					Nucleotide range of extended sequence (bp #)	Database nucleotide range (bp #)
14	IMX2_55.EXT	332 – 744	Homology to R. norvegicus C-reactive protein mRNA, complete cds. (M83176.1)	65%	119 – 253 303 – 567	115 – 253 299 – 563
			Murine mRNA for C-reactive protein (CRP) (X17496.1)	56% 62%	4 – 118 194 – 567	80 – 194 277 – 650
			Mus musculus C-reactive protein, petaxin related (Crp), mRNA (NM_007768.1)	56% 62%	4 – 118 194 – 567	80 – 194 277 – 650
16	IMX2_57.EXT	283 – 408	Homology to H. sapiens mRNA for chymotrypsin-like protease CTRL-1 (X71877.1)	84%	7 – 363	464 – 820
			Homology to Homo sapiens chymotrypsin-like (CTRL) mRNA (NM_001907.1)	84%	7 – 363	464 – 820
			Homology to H. sapiens genes for proteasome-like subunit (MECL-1), chymotrypsin-like protease (CTRL-1) and protein serine kinase (PSK-H1) last exon	83% 83% 84%	7 – 55 52 – 187 179 – 363	8641 – 8689 8773 – 8908 8990 – 9174
18	IMX2_61.EXT	204 – 425	Homology not determined			

EXAMPLE 4

Extended Sequence Clone for IMX 2_4

The clone was isolated from a D4/D8 DSS colon library using the Clonecapture procedure from Edge Biosystems. The library was constructed by Edge Biosystems for Immunex from RNA isolated from D4 and D8 inflamed colons from DSS treated C57BL/6 mice. The clone is in an Edge vector pEAK12.

IMX2_04 DST (SEQ ID NO:1) matches extended sequence for IMX2_04 from base 688 to 957 (SEQ ID NO:2) (See Table 7, above). There is an open reading frame starting at base 1 that may encode a partial gene product. The amino acid sequence of the encoded partial protein is given in SEQ ID NO:129.

When the extended sequence was compared to NCBI nr database w/BLAST, linear segments of the nucleotide sequence for bases 1-936 show about 70% identity with segments of the nucleotide sequence of a cosmid clone from human chromosome 19, AC007565.1 (See Table 7, above). This indicates that the human homologue of IMX2_4 is, at least in part, found on the cosmid. However, several exons predicted on the cosmid (by GRAIL program or by some homology to mouse ESTs) are skipped by the linear sequence of IMX2-4-E10.seq. It appears that one predicted, yet skipped exon, is real in that a perfect match is found in the DERWENT database. The Derwent entry is for a "secreted" molecule with the protein fragment in the Derwent protein entry being a signal peptide containing amino acid sequence not found on the cosmid (or the nucleotide Derwent entry). However, the Derwent nucleotide entry also has a match to another more 5' segment of the cosmid which does show a match with the 5' end of IMX2-4-E10.seq at the 75% identity level. These forms may represent different splice variants of a secreted protein.

EXAMPLE 5

Extended Sequence Clone for IMX 2_36

The IMX2_36.EXT sequence information was derived from the clone IMX2_36pT7T3-2.seq which is an EST clone derived from FVB/N mouse proximal colon obtained from IMAGE consortium. The accession number for the EST is AA529850, and the

clone id is IMAGE:921608. The EST is in the pT7T3 vector. The IMX2_36 DST (SEQ ID NO:6) matches the extended sequence for IMX2_36.EXT (SEQ ID NO:7) from base 293 to 427 (See Table 7, above). Blast of the EST to GenBank gives hits to 'mouse KAR (killer activating receptor)' and 'DAP12 protein' (See Table 7, above).

EXAMPLE 6

Extended Sequence Clone for IMX 2_43

The clone was isolated from a D4/D8 DSS colon library using the Clonecapture procedure from Edge Biosystems. The library was constructed by Edge Biosystems for Immunex from RNA isolated from pooled D4 and D8 inflamed colons from DSS treated C57BL/6 mice. The clone is in the Edge vector pEAK12. The IMX2_43 DST (SEQ ID NO:8) matches the IMX2_43 extended sequence (SEQ ID NO:9) from base 1079 to 1298 (See Table 7, above).

EXAMPLE 7

Extended Sequence Clone for IMX 2_46

The IMX2_46.EXT sequence information was derived from the two clones IMX2_46pT7T3-6.seq and IMX2_46pT7T3-7.seq, which are EST clones in the pT7T3 vector that were obtained from IMAGE consortium. IMX2_46pT7T3-6.seq (accession number AA290194, clone id IMAGE:750847) was derived from C57BL/6 mouse lymph node. IMX2_46pT7T3-7.seq (accession number AA174968, clone id IMAGE: 617717) was derived from C57BL/6 mouse spleen. The IMX2_46 DST (SEQ ID NO:10) aligns with the IMX2_46.EXT extended sequence (SEQ ID NO:11) from base 157 to 561 (See Table 7, above).

Blast of ESTs against GenBank gives hits to human TOSO: regulator of fas-induced apoptosis (See Table 7, above).

EXAMPLE 8

Extended Sequence Clone for IMX 2_55

The IMX2_55.EXT extended sequence information was derived from the two clones IMX2_55pT7T3-8.seq and IMX2_55pT7T3-24.seq, which are EST clones that were obtained

from IMAGE consortium. IMX2_55pT7T3-8.seq (accession number AA823573, clone id IMAGE: 1079189) was derived from FVB irradiated mouse colon. IMX2_55pT7T3-24.seq (accession number AA690372, clone id IMAGE:1164692) was derived from from FVB mouse proximal colon.

The accession number for the EST for IMX2_55pT7T3-8.seq is AA823573, and the clone id is IMAGE:1079189. The accession number for the EST for IMX2_55pT7T3-24.seq is AA690372, and the clone id is IMAGE:1164692. The both ESTs are in the pT7T3 vector and were obtained from IMAGE consortium. IMX2_55pT7T3-8.seq clone was derived from FVB irradiated mouse colon and IMX2_55pT7T3-24.seq clone was derived from FVB mouse proximal colon. The IMX2_55 DST (SEQ ID NO:13) aligns bases of 322 to 744 with the extended sequence IMX2_55.EXT (SEQ ID NO:14) (See Table 7, above).

Blast of the extended IMX2_55 sequence to GenBank disclosess ESTs with homology to C reactive protein (See Table 7, above).

EXAMPLE 9

Extended Sequence Clone for IMX 2_57

PCR primers were designed for IMX2_57 based on sequence information obtained from the sequences of two EST clones that were commercially unavailable. The EST clones used for deriving sequence information were (1) an EST with accession number AA240177 and clone id IMAGE:679264 derived from mouse liver and (2) an EST with accession number AV005227 and clone id 0910001G08 derived from C57BL/6 spleen.

The primers used were forward primer IMX2_57-FP and reverse primer IMX2_57-RP which prime off the sequence obtained from electronically assembling the DST and the EST AV005227. The PCR reaction was performed on cDNA derived from RNA prepared from C57BL/6 mouse colon. The PCR product was cloned into the pGEM vector resulting in the clone IMX2_57PCR1. The IMX2_57 DST, (SEQ ID NO:15) aligns with bases 283 to 408 of the extended sequence IMX2_57.EXT (SEQ ID NO:16) (See Table 7, above).

Blast of assembled ESTs to Genbank gives hits to 'human chymotrypsin-like (CTRL) mRNA' and 'human proteosome-like subunit (MECL-1), chymotrypsin-like protease (CTRL-1) and protein-serine kinase (PSK-H1) last exon (See Table 7, above).

EXAMPLE 10

Extended Sequence Clone for IMX 2_61

The IMX2_61 extended sequence information was derived from the EST clone IMX2_61pBS-47.seq, an EST clone that was obtained from IMAGE consortium. The accession number for the EST for IMX2_61pBS-47.seq is AA981092, and the clone ID is IMAGE:1279287. The IMX2_61pBS-47.seq clone was derived from WEHI3 mouse macrophage cells. The IMX2_61 DST (SEQ ID NO:17) aligns with bases 204 to 425 of the IMX2_61 extended sequence (SEQ ID NO:18) (See Table 7, above).

EXAMPLE 11

Extended Sequence Clone for IMX 2_17: Crypt-ductin alpha scavenger receptor (CRP-ductin)

TOGA analysis indicated that IMX2_17 (SEQ ID NO:3) corresponds to *Mus musculus* CRP-ductin-alpha mRNA, accession number U37438 (Table 3). CRP-ductin localizes to the apical portion of crypt cells in the small intestine. In the colon, it is seen predominantly in surface epithelial cells (EC). It is also seen in the apical portion of the EC lining pancreatic and larger hepatic ducts. The CRP-ductin-alpha sequence predicts a mosaic protein with a short cytoplasmic region, a transmembrane domain and a large extracellular region composed of many repeats. The extracellular region contains 21 potential N-glycosylation sites in the C-terminal half of the protein. There are also two potential phosphorylation sites in the cytoplasmic domain.

Forward primer mCRP.5179-FP and reverse primer mCRP.6106-RP were used to PCR a 936 bp product from a cDNA template derived from RNA isolated from C57BL/6 colon. The PCR product was subcloned into the pGEM vector and the sequence was verified before being used as a probe for Northern blot analysis.

Figure 6 presents the results of Northern blot analysis of clone IMX 2_17, SEQ ID NO: 3, where an agarose gel containing poly A enriched mRNA from the experimental samples from and size standards was blotted after electrophoresis, probed with a labeled probe corresponding to U37438 bases 5179-6106, imaged using a phosphorimager and quantified. Figure 6A shows the results from C57BL/6 mice with DSS colitis 0, 4, 8, or 12 days after treatment, C57BL/6 mice with αCD3 ileitis 0, 6, 30 or 72 hours after treatment, as well as samples from large intestines from FVB, mdr knock-out mice without colitis and mdr knock-out mice with colitis. Figure 6B shows the results from Balb/c mice with DSS colitis 0, 4, 8, or 12 days after treatment, Balb/c mice treated with 0%, 5% and 8% DSS, and Balb/c mice with αCD3 ileitis 0, 6, 30 or 72 hours after treatment. The predicted transcript size for IMX2_17 CRP-ductin is 6.6 Kb; the actual transcript size found in this study was approximately 6.5 Kb.

Northern blots were performed on mRNA samples from several models of IBD. Analysis of samples from C57BL/6 mice with DSS colitis showed high levels of constitutive expression in the large intestine, maximal at day 8 of DSS colitis (Figure 6A), consistent with the results of the initial TOGA analysis (Table 1). Analysis of samples from Balb/c mice with DSS colitis showed similar results (Figure 6B).

Analysis of samples from C57BL/6 mice with αCD3 ileitis showed low levels of constitutive expression in the small intestine that increases early in inflammation (Figure 6A) reaching a maximum at six hours and declining thereafter. Analysis of samples from Balb/c mice with αCD3 ileitis showed similar, though less intense, results (Figure 6B).

Constitutive expression was seen in FVB large intestine samples (Figure 6A). Increased expression in healthy mdr knock-out mice with no signs of colitis, with little further increase in expression in mdr knock-out mice with active colitis (Figure 6A).

The expression shown on the Northern blot of Figure 6 was quantified and normalized to the amount of G3PDH in each lane. The normalized results are shown in Table 8, below.

Table 8 Quantitation of Figure 6		
IBD Model	**C57BL/6**	**BALB/c**
DSS induced colitis		
Day 0	27740	11455
Day 4	30853	21103
Day 8	43857	28838
Day 12	22622	26248
DSS Concentration Effect		
0 %	Not Done	11832
5 %	Not Done	21278
8 %	Not Done	12717
Anti CD3 Induced Ileitis		
0 Hours	2519	2491
6 Hours	14088	7870
30 Hours	11144	1287
72 Hours	9000	2238
Constituitive Expression		
FVB	**Mdr Knock Out**	**Mdr Knock Out + Colitis**
14427	29166	15988

EXAMPLE 12

Extended Sequence Clone for IMX 2_22

IMX2_22 Hematopoietic Progenitor Kinase HPK1

TOGA analysis indicated that IMX2_22 corresponds to *Mus musculus* mRNA for serine/threonine kinase, accession number Y09010 (Table 3). HPK1 is a hematopoietic protein kinase activating the SAPK/JNK pathway.

Forward primer mHPK1.1640-FP and reverse primer mHPK1.2420-RP were used to PCR a 780 bp product from a cDNA template was derived from RNA isolated from C57BL/6 colon. The PCR product was subcloned into the pGEM vector and was sequence verified before being used as a probe for Northern blot analysis. The predicted transcript size for IMX2_22 HPK1 was 2.7 Kb and published transcript sizes are 2.8 Kb and 3.6 Kb. The actual transcript size found in this study was 2.8Kb.

Figure 7 presents the results of Northern blot analysis of clone IMX 2_22, SEQ ID NO: 4, where an agarose gel containing poly A enriched mRNA from the experimental samples and size standards was blotted after electrophoresis, probed with a labeled probe corresponding to bases 1640-2420 of Y09010, imaged using a phosphorimager and

quantified. Figure 7A shows the results from C57BL/6 mice with DSS colitis 0, 4, 8, or 12 days after treatment, C57BL/6 mice with αCD3 ileitis 0, 6, 30 or 72 hours after treatment, as well as samples from large intestines from FVB, mdr knock-out mice without colitis, mdr knock-out mice with colitis and C57BL/6 spleen. Figure 7B shows the results from Balb/c mice with DSS colitis 0, 4, 8, or 12 days after treatment, Balb/c mice with αCD3 ileitis 0, 6, 30 or 72 hours after treatment and C57BL/6 normal lymphoid tissue samples (MLM, PP, spleen and thymus).

Northern blots were performed on mRNA samples from several models of IBD. Analysis of samples from C57BL/6 mice with DSS colitis showed low levels of constitutive expression in the large intestine, maximal at day 12 of DSS colitis (Figure 7A), not very consistent with the results of the initial TOGA analysis (Table 1). Analysis of samples from Balb/c mice with DSS colitis showed similar results with a maximum at day 8 (Figure 7B).

Analysis of samples from C57BL/6 mice with αCD3 ileitis showed constitutive expression in the small intestine that decreases at six hours (Figure 7A). Analysis of samples from Balb/c mice with αCD3 ileitis showed similar results (Figure 7B).

Increased expression in mdr knock-out mice with active colitis (Figure 7A), in contrast to little change in corresponding TOGA analysis (data not shown).

Strong expression was found in lymphoid tissues in MLN, PP, thymus, especially the spleen (Figure 7A & B).

The expression shown on the Northern blot of Figure 7 was quantified and normalized to the amount of G3PDH in each lane. The normalized results are shown in Table 9, below.

Table 9
Quantitation of Figure 7

IBD Model	C57BL/6	BALB/c
DSS induced colitis		
Day 0	-50	271
Day 4	64	143
Day 8	82	502
Day 12	419	341

Anti CD3 Induced Ileitis		
0 Hours	614	369
6 Hours	133	-47
30 Hours	465	341
72 Hours	168	149

Constituitive Expression		
FVB	Mdr Knock Out	Mdr Knock Out + Colitis
59	146	739

Normal Tissue			
MLN	PP	Spleen	Thymus
10008	4883	33260	8826

EXAMPLE 13

Extended Sequence Clone for IMX 2_28

Down-Regulated in Adenoma protein (DRA)

TOGA analysis indicated that IMX2_28 corresponds to *Mus musculus* DRA down-regulated in adenoma protein, accession number AF136751 (Table 3).

Forward primer Dra.1551-FP and reverse primer Dra.2390-RP were used to PCR an 840 bp product from a cDNA template derived from RNA isolated from C57BL/6 colon. The PCR product was subcloned into the pGEM vector and was sequence verified before being used as a probe for Northern blot analysis. The predicted transcript size for IMX2_28 DRA is 2.6 Kb; the actual transcript size found in this study was approximately 3 Kb.

Figure 8 presents the results of Northern blot analysis of clone IMX 2_28, SEQ ID NO: 5, where an agarose gel containing poly A enriched mRNA from the experimental samples and size standards was blotted after electrophoresis, probed with a labeled probe corresponding to bases 1551-2390 of AF136751, imaged using a phosphorimager and quantified. Figure 8A shows the results from C57BL/6 mice with DSS colitis 0, 4, 8, or 12 days after treatment, C57BL/6 mice with αCD3 ileitis 0, 6, 30 or 72 hours after treatment, as well as samples from large intestines from FVB, mdr knock-out mice without colitis and mdr knock-out mice with colitis. Figure 8B shows the results from Balb/c mice with DSS colitis 0, 4, 8, or 12 days after treatment, Balb/c mice treated with 0%, 5% and 8% DSS, and Balb/c mice with αCD3 ileitis 0, 6, 30 or 72 hours after treatment.

Northern blots were performed on mRNA samples from several models of IBD. Analysis of samples from C57BL/6 mice with DSS colitis showed high levels of constitutive expression in the large intestine that decrease with inflammation, minimal at day 8 of DSS colitis (Figure 8A), consistent with the results of the initial TOGA analysis (Table 1). Analysis of samples from Balb/c mice with DSS colitis showed the same pattern (Figure 8B).

Analysis of samples from C57BL/6 mice with αCD3 ileitis showed levels of constitutive expression in the small intestine lower than that seen in the large intestine with DSS colitis. Expression decreases with inflammation rapidly by 6 hours and remains low (Figure 8A). Analysis of samples from Balb/c mice with αCD3 ileitis showed the same pattern (Figure 8B). TOGA analysis in this model (data not shown) did not show a corresponding peak.

Constitutive expression was seen in FVB large intestine samples (Figure 8A). Constitutive expression was seen in healthy mdr knock-out mice with no signs of colitis, with a dramatic decrease in expression seen in mdr knock-out mice with active colitis (Figure 8A). TOGA analysis in this model (data not shown) did not show a corresponding peak.

The expression shown on the Northern blot of Figure 8 was quantified and normalized to the amount of G3PDH in each lane. The normalized results are shown in Table 10, below.

Table 10 Quantitation of Figure 8		
IBD Model	**C57BL/6**	**BALB/c**
DSS induced colitis		
Day 0	25541	34441
Day 4	20316	35270
Day 8	7569	22711
Day 12	10218	22481
DSS Concentration Effect		
0 %	Not Done	22463
5 %	Not Done	19861
8 %	Not Done	17548
Anti CD3 Induced Ileitis		
0 Hours	5937	3534
6 Hours	2211	1275
30 Hours	2192	2051
72 Hours	5300	4142
Constituitive Expression		
FVB	**Mdr Knock Out**	**Mdr Knock Out + Colitis**
55232	28854	5046

EXAMPLE 14

Extended Sequence Clone for IMX2_33

Secretory Leukocyte Protease Inhibitor (SLPI)

TOGA analysis indicated that IMX2_33 corresponds to *Mus musculus* secretory leukocyte protease inhibitor, accession number U73004 (Table 3). Secretory leukocyte protease inhibitor is an epithelial cell and macrophage derived inhibitor of leukocyte serine proteases. SLPI expression is suppressed by gamma-IFN. SLPI is an LPS induced gamma-IFN suppressible phagocyte product that serves to inhibit LPS responses.

Forward primer mSLPI.447-FP and reverse primer mSLPI.800-RP were used to PCR a 350 bp product from a cDNA template was derived from RNA isolated from C57BL/6 colon. The PCR product was subcloned into the pGEM vector and the sequence was verified before being used as a probe for Northern blot analysis. The probe used corresponded to bases 447-800 of U73004. The predicted transcript size for IMX2_33 SLPI is 1.1 Kb; the actual transcript size found in this study was 1.1 Kb.

Figure 9 presents the results of Northern blot analysis of clone IMX 2_33, SEQ ID NO:21, where an agarose gel containing poly A enriched mRNA from the experimental samples and size standards was blotted after electrophoresis, probed, imaged using a phosphorimager and quantified. Figure 9A shows the results from C57BL/6 mice with DSS colitis 0, 4, 8, or 12 days after treatment, C57BL/6 mice with αCD3 ileitis 0, 6, 30 or 72 hours after treatment, as well as samples from large intestines from FVB, mdr knock-out mice without colitis and mdr knock-out mice with colitis. Figure 9B shows the results from Balb/c mice with DSS colitis 0, 4, 8, or 12 days after treatment, Balb/c mice treated with 0%, 5% and 8% DSS, and Balb/c mice with αCD3 ileitis (0, 6, 30, 72 hours).

Northern blots were performed on mRNA samples from several models of IBD. Analysis of samples from C57BL/6 mice with DSS colitis showed minimal constitutive expression in the large intestine that increases significantly with inflammation, maximal at day 8 of DSS colitis and still high at day 12 (Figure 9A), consistent with the results of the initial TOGA analysis (Table 1). Analysis of samples from Balb/c mice with DSS colitis showed the same pattern (Figure 9B).

Analysis of samples from C57BL/6 mice with αCD3 ileitis showed low to moderate levels of constitutive expression in the small intestine with less regulation than seen in the large intestine with DSS colitis. There was a significant increase with inflammation maximal at 30 hours, then decreasing (Figure 9A). Analysis of samples from Balb/c mice with αCD3 ileitis showed the same pattern (Figure 9B). The pattern was consistent with that seen in TOGA analysis in this model (data not shown).

Minimal constitutive expression was seen in healthy mdr knock-out mice with no signs of colitis, with an increase in expression seen in mdr knock-out mice with active colitis (Figure 9A). The pattern was consistent with that seen in TOGA analysis in this model (data not shown).

The expression shown on the Northern blot of Figure 9 was quantified and normalized to the amount of G3PDH in each lane. The normalized results are shown in Table 11, below.

Table 11 Quantitation of Figure 9		
IBD Model	**C57BL/6**	**BALB/c**
DSS induced colitis		
Day 0	78	197
Day 4	1580	2237
Day 8	7491	8170
Day 12	6313	4390
DSS Concentration Effect		
0 %	Not Done	-54
5 %	Not Done	3201
8 %	Not Done	-688
Anti CD3 Induced Ileitis		
0 Hours	658	180
6 Hours	832	1100
30 Hours	1465	2572
72 Hours	526	1245
Constituitive Expression		
FVB	**Mdr Knock Out**	**Mdr Knock Out + Colitis**
965	14505	30085

EXAMPLE 15

Extended Sequence Clone for IMX 2_48

IMX2_48 macrophage inflammatory protein 2 (MIP2)

TOGA analysis indicated that IMX2_48 corresponds to *Mus musculus* MIP-2, macrophage inflammatory protein 2, accession number X53798 (Table 3).

Forward primer MIP2.61-FP and reverse primer MIP2.345-RP were used to PCR a 284 bp product from a cDNA template was derived from RNA isolated from C57BL/6 colon. The PCR product was subcloned into the pGEM vector and the sequence was verified. The predicted transcript size for IMX2_48 MIP2 is 1.1 Kb.